E R R A T A

Technical Report No. 1 (USWB Contract Cwb-10704)

REGRESSION ESTIMATION OF EVENT PROBABILITIES

April 1964

Page 4 - Line 4 - The word "divided" should be "dividing."

Page 6 - Title for Experiment 8 should read "Combination of Selected Predictors from a Network and Single-Station Booleans."

Note: This would also change the title on Page 129.

Page 26 - Line 12 - Insert "City" after Atlantic.

Page 69 - Line 4 - The word "of" should be "by."

- Line 7 (from the bottom) - The word "dependent" should be

U.S. Weather Bureau
Contract Cwb-10704
Technical Report No. 1

AD 602037

REGRESSION ESTIMATION OF EVENT PROBABILITIES

Robert G. Miller

April 1964
7411-121

This report, prepared by The Travelers Research Center, Inc., does not necessarily reflect official policy or views of the sponsoring agency.

THE TRAVELERS RESEARCH CENTER, INC.
250 Constitution Plaza — Hartford, Connecticut 06103

ABSTRACT

MDA (multiple discriminant analysis) has been found useful in predicting the probability of operationally critical airfield ceiling and visibility conditions. One part of the discriminant procedure requires at the outset a rather lengthy computation by machine to estimate these probabilities. While the operational use of the method takes much less effort, the time needed initially on a large scale computer may be inordinate when many elements, time periods, and geographical locations are being considered. This technical report discusses a prediction technique, REEP (regression estimation of event probabilities), which resulted from attempts to make MDA more efficient. Comparative studies between REEP and MDA are included as well as other experimentations performed using the REEP procedure. Recommendations for future work on the technique are also given.

TABLE OF CONTENTS

		Page
1.0	Introduction	1
2.0	The REEP Procedure	2
3.0	REEP Predictor Selection	5
4.0	Experimentation	6
5.0	Discussion	145
6.0	Future Work	150
7.0	Acknowledgements	152
8.0	Bibliography	153

1.0 INTRODUCTION

The objective of this task was to develop, in the MDA (multiple discriminant analysis) framework, a method which would produce medium range ceiling and visibility forecasts efficiently. The original plan was to investigate the use of a multivariate statistical procedure called canonical correlation--discriminant analysis being a special case of canonical correlation. By incorporating a feature which screened predictors in canonical correlation, it was expected that more than one predictand could be processed simultaneously with a resulting saving in computer time.

Work on screening canonical correlation was proceeding satisfactorily until it was found that a technique subsequently called REEP (regression estimation of event probabilities) possessed even more desirable computational features. It was later discovered that the REEP technique had appeared in the literature in 1955, Lund [1], (see also Lund [2]). In addition, work related to this approach has been published by Cox [3, 4] and by Warner [5].

This technical report consists of 1) a discussion of the REEP technique, 2) the results of a comparative study between REEP and MDA, 3) the results from a series of experiments investigating various aspects of the REEP procedure, 4) the results from experiments performed to study the effect of modifying the form of the predictor variables by means of Boolean expressions, and 5) discussion and recommendations for future study.

2.0 THE REEP PROCEDURE

Consider the problem of estimating the probability distribution over a set of G mutually exclusive and exhaustive groups where the groups are defined for a specific predictand. The statistical technique of discriminant analysis may be used to estimate probability distributions of this type [6]. The suggested approach for estimating the group probabilities, when the underlying density functions are unknown, is that of Fix and Hodges. This approach requires the computation of "distances" in the discriminant space between a new observation and <u>each</u> observation of the developmental sample. The set of k close neighbors is then used to estimate the group probabilities.

The REEP procedure may be thought of as an alternative to the Fix-Hodges method employed in MDA. Or, more appropriately, it can be viewed as a self-contained statistical prediction method. The following is a description of the REEP procedure:

Perform a series of multiple regressions on G zero-one dependent variables, $Y_1,\ldots,Y_G$, where each dependent variable is associated with one of G predictand groups. The independent variables, $X_1,\ldots,X_r$, are taken to be the same throughout each of the G regressions. The object of this series of regressions is to obtain the least squares estimate of the A's in the following equations:

$$E(Y_1|\underline{X}) = \sum_{s=0}^{r} A_{1s}X_s$$

$$E(Y_2|\underline{X}) = \sum_{s=0}^{r} A_{2s}X_s$$

$$\vdots$$

$$E(Y_G|\underline{X}) = \sum_{s=0}^{r} A_{Gs}X_s$$

where $X_0 \equiv 1$

All of the conditional distributions, $f(Y_g|\underline{X})$, $g=1,\ldots,G$, are Bernoulli distributions (zero-one) and on a single trial the expectation is equal to the probability that $Y_g = 1$. Thus the regression functions give least squares estimates of the group probabilities, $\hat{P}_g$ $(g=1,\ldots,G)$.

These estimates possess certain desirable and certain undesirable properties:

Desirable

1) The estimates add to unity, i.e., $\sum_{g=1}^{G} \hat{P}_g = 1$.

2) The estimates essentially minimize the Brier-Allen $\overline{P}$ score,

i.e., $\sum_{i=1}^{N} \sum_{g=1}^{G} (Y_{ig} - \hat{P}_{ig})^2$.

Undesirable

1) The estimates are not bounded by 0 and 1.

An intuitively reasonable approach and the one used to remedy this undesirable property is to renormalize the estimates by: a) setting all negatives equal to zero, b) setting all estimates greater than one equal to one, and c) divided each by the resulting overall sum.

3.0 REEP PREDICTOR SELECTION

A predictor selection procedure has been incorporated into the REEP procedure similar to that used in regression and MDA [6], namely:

Out of a set of P possible predictors $X^{(1)}$ is selected such that in the developmental sample it contributes most significantly with regard to one of the G predictands. That is, out of the G possible F distribution statistics the predictor $X^{(1)}$ has one of these computed statistics larger than any one of the G computed for the remaining P-1 predictors.

In a similar manner, a second predictor $X^{(2)}$ is selected from the P-1 remaining predictors when considered in conjunction with $X^{(1)}$. Selection continues until r predictors $X^{(1)}, X^{(2)}, \ldots, X^{(r)}$ have been chosen; r is determined such that $X^{(r+1)}$ is not satisfactorily significant. The significance criterion has been designed similar to that used in screening regression and MDA (see page 13, reference [6]). Namely, compare the G possible computed F statistics, at any of the r selection stages, with a critical value of F_{α}, where the size α of the test is $\frac{1}{20\left[P - (S-1)\right]}$. The value S denotes the selection stage number or equivalently the number of predictors thus far selected. Tests must still be performed to confirm its applicability, however.

4.0 EXPERIMENTATION

A number of experiments were performed to test certain aspects of the REEP procedure. The complete list of experiment titles is given below.

	Title	Page
Experiment 1 -	Comparison Between MDA and REEP (Atlantic City Airport).	7
Experiment 1A -	Comparison Between MDA and REEP (Offutt AFB, Nebraska).	25
Experiment 2 -	Standard Dummies versus Cumulative Dummies	44
Experiment 3 -	Nullifying the Effect of High Ceilings in Predictor Selection.	69
Experiment 4 -	Ceiling Prediction Using Persistence with One and Two Hour Lagged Persistence.	78
Experiment 5 -	Multiple Predictands Processed Simultaneously.	89
Experiment 6 -	Effect of Restricting Selection to Predictors with Certain Minimum Occurrence Frequencies.	96
Experiment 7 -	Single-Station Predictors: Raw Dummies and Boolean Combinations.	106
Experiment 8 -	Combination of Selected Predictors from Network and Single-Station Boolean.	129

EXPERIMENT 1. Comparison Between MDA and REEP (Atlantic City Airport)

Purpose

To compare the results of short range ceiling and visibility predictions using the MDA technique and the REEP technique. In the event REEP compares favorably with MDA it might, because of its computational efficiency, be used in place of MDA for forecasting ceiling and visibilities.

Procedure

A test of the REEP technique was made by comparing independent sample predictions of ceiling and visibility using REEP and MDA. Each technique was allowed to select predictors and generate probability forecasts independently. Predictions were made for two stations, Atlantic City Airport, New Jersey and Offutt AFB, Nebraska. (See Experiment 1A for Offutt.)

Predictands

The five operationally significant categories of ceiling and visibility at Atlantic City Airport, New Jersey for the forecast intervals of three and seven hours define the four predictands for which the techniques are compared (see Table 1-1). The specifications for each predictand are given in Table 1-2.

Predictors

Seven meteorological elements at 13 predictor stations (see Figure 1-1) plus the time of day (TOD) and day of year (DOY) variables were transformed into dummy (zero-one) variables (see Tables 1-3 and 1-4). This set of 423 dummy predictors was used as possible predictors for all four predictands.

Data Sample

Standard hourly airways surface observations covering the 10 year period, 1 January 1949 to 31 December 1958, furnished the data for this study. There are 87,672 hours in the 10 years from 1949 through 1958. This is too many to use efficiently; besides not all hours are needed because of the serial correlation present in meteorological variables. The amount of data was reduced by a three-stage process:

(a) Any variable not observed for a substantial portion of the 87,672 hours was eliminated. A variable was also eliminated if the frequency distribution of its error vector exhibited a bias with regard to hour of day or month of year.

(b) Those hours were eliminated for which any one of the variables was missing or failed to pass a gross-error check.

(c) A random selection of the remaining hours was made to select 10,000 cases.

Results

The four sets of predictors selected by MDA and REEP respectively are given in Tables 1-5a to 1-5d as well as the REEP regression equation coefficients for each group of the predictand. Contingency tables (see Tables 1-6a to 1-6d) show the forecast-versus-observed frequencies for persistence, MDA, and REEP. The number of correct forecasts (hits) and the Brier and Allen $\bar{P}$ scores as well as the individual group contributions to the overall $\bar{P}$ score are given below each contingency table. The number of hits and the overall $\bar{P}$ scores are summarized in Table 1-7 by predictand.

Conclusions

See Experiment 1A.

Table 1-1

Definition of Predictand Groups for Atlantic City Municipal Airport

Group Number	Ceiling Intervals
1	0 ft ≤ Ceiling < 200 ft
2	200 ft ≤ Ceiling < 500 ft
3	500 ft ≤ Ceiling < 1000 ft
4	1000 ft ≤ Ceiling < 3000 ft
5	3000 ft ≤ Ceiling

Group Number	Visibility Intervals
1	0 mi ≤ Visibility < 1/2 mi
2	1/2 mi ≤ Visibility < 1 mi
3	1 mi ≤ Visibility < 2 mi
4	2 mi ≤ Visibility < 3 mi
5	3 mi ≤ Visibility

Table 1-2

Specification of Predictands
Atlantic City Municipal Airport

	Predictand 1	2	3	4
Predictand variable	CIG	CIG	VIS	VIS
Forecast interval--hours (H)	3	7	3	7
Number of Groups (G)	5	5	5	5
Observations in Group 1 (n_1)	100	135	153	167
Observations in Group 2 (n_2)	281	272	125	130
Observations in Group 3 (n_3)	385	420	200	230
Observations in Group 4 (n_4)	660	619	286	283
Observations in Group 5 (n_5)	6530	6510	7192	7146
Total dependent sample size (N)	7956	7956	7956	7956
Number of independent observations in the dependent sample (n)	2652	2652	2652	2652
Total independent sample size (M)	2044	2044	2044	2044
Number of available predictors (P)	423	423	423	423

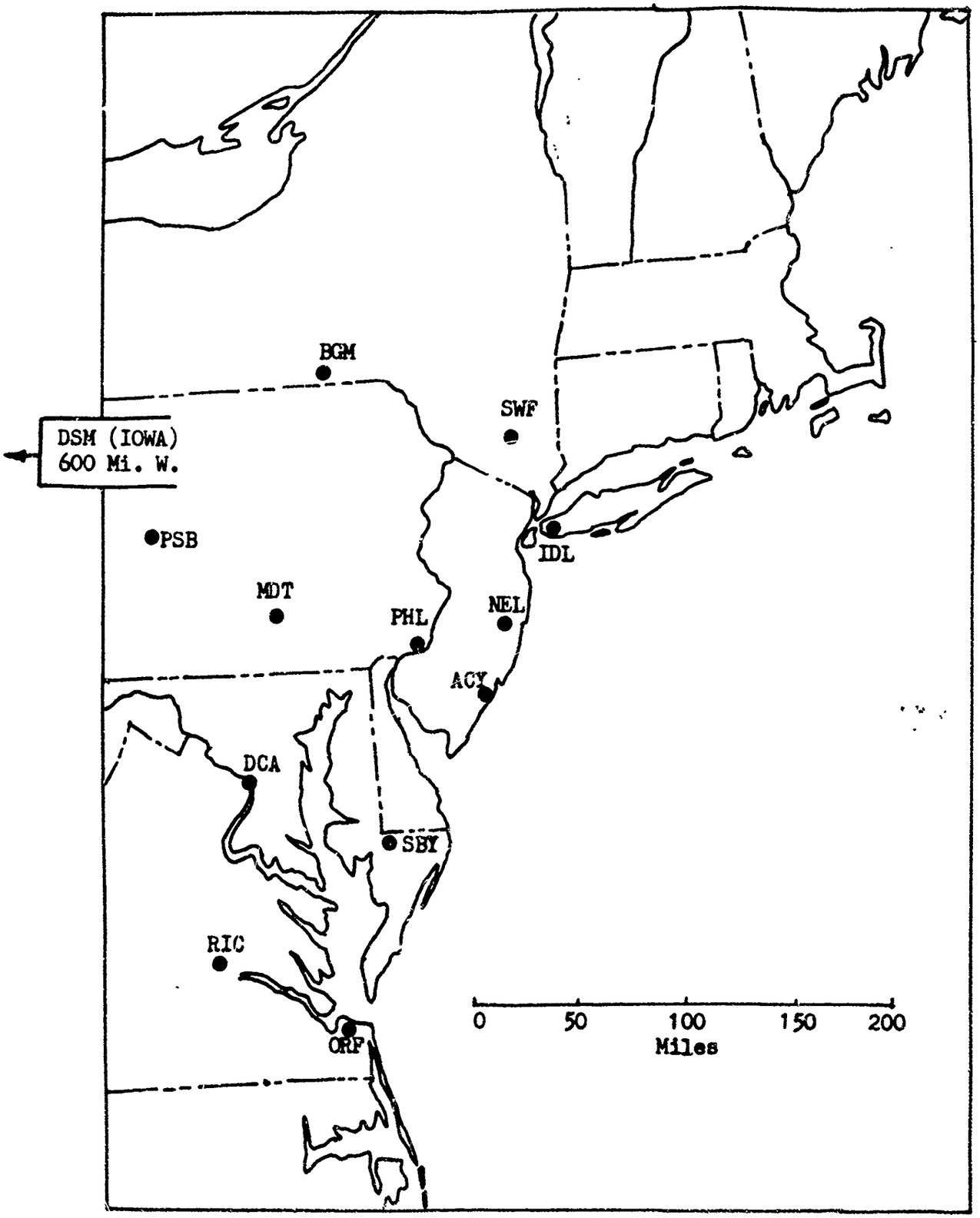

Figure 1-1. Thirteen Station Network for Atlantic City Airport

Table 1-3

Number of Possible Predictors for the
3- and 7-hr Ceiling and Visibility at ACY

	Station													
Elem.	ACY	NEL	PHL	SBY	MDT	DCA	ORF	PSB	BGM	SWF	RIC	IDL	DSM	Total
TOD	5	-	-	-	-	-	-	-	-	-	-	-	-	5
DOY	1	-	-	-	-	-	-	-	-	-	-	-	-	1
CIG	15	5	5	5	5	5	5	5	5	5	5	5	-	70
VIS	15	5	5	5	5	5	5	5	5	5	5	5	-	70
WND	5	5	5	5	5	5	5	5	5	5	5	5	5	65
RLH	3	3	3	3	3	3	3	3	3	3	3	3	3	39
TCA	2	2	2	2	2	2	2	2	2	2	2	2	2	26
WEA	12	12	12	12	12	12	-	12	12	12	12	12	12	144
SCL	3	-	-	-	-	-	-	-	-	-	-	-	-	3
												Total		423

Table 1-4

Specifications and Limits of the Zero-One (Dummy) Variables

Meteorological Element	Number of Dummies	Limits of Each Category
		Time of Day:
TOD	5	1. From 0130 EST to 1330 EST
		2. From 0130 EST to 0730 EST
		3. From 0730 EST to 1330 EST
		4. From 1330 EST to 1930 EST
		5. From 1930 EST to 0130 EST
		Day of Year:
DOY	1	1. From 151 (31 May) to 300 (27 Oct)
		Wind:
WND	5	1. Calm to 3 Knots, any direction
		2. From 23° to 67°, 3 Knots < Speed
		3. From 68° to 202°, 3 Knots < Speed
		4. From 203° to 247°, 3 Knots < Speed
		5. From 248° to 22°, 3 Knots < Speed
		Relative Humidity:
RLH	3	1. From 75 per cent to 85 per cent
		2. From 86 per cent to 93 per cent
		3. From 94 per cent to 100 per cent
		Total Cloud Amount:
TCA	2	1. From 5/10 to 9/10 (incl.)
		2. More than 9/10
		Ceiling Height, ft:
CIG	5	1. $0 \leq CIG < 200$
		2. $200 \leq CIG < 500$
		3. $500 \leq CIG < 1000$
		4. $1000 \leq CIG < 3000$
		5. $3000 \leq CIG$

Table 1-4 - Continued

Meteorological Element	Number of Dummies		Limits of Each Category
			Ceiling Height*, ft:
CIG	10	(1)	CIG = 100
		(2)	CIG = 200
		(3)	CIG = 400
		(4)	CIG = 500
		(5)	CIG = 900
		(6)	1500 ≤ CIG < 2000
		(7)	2500 ≤ CIG < 3000
		(8)	3000 ≤ CIG < 5000
		(9)	5000 ≤ CIG < 10,000
		(10)	CIG = Unlimited
			Visibility, mi:
VIS	5	1.	0 ≤ VIS < 1/2
		2.	1/2 ≤ VIS < 1
		3.	1 ≤ VIS < 2
		4.	2 ≤ VIS < 3
		5.	3 ≤ VIS
			Visibility*, mi:
VIS	10	(1)	1/4 ≤ VIS < 1/2
		(2)	1/2 ≤ VIS < 3/4
		(3)	3/4 ≤ VIS < 1
		(4)	1 ≤ VIS < 1-1/2
		(5)	1-1/2 ≤ VIS < 2
		(6)	2 ≤ VIS < 3
		(7)	3 ≤ VIS < 4
		(8)	4 ≤ VIS < 5
		(9)	5 ≤ VIS < 6
		(10)	15 ≤ VIS

*Parentheses are used here to help distinguish the two forms of ceiling and visibility dummies.

Table 1-4 - Continued

Meteorological Element	Number of Dummies	Limits of Each Category	
		Weather Element:	Teletype Code:
WEA	12	1. No Weather	
		2. Rain	R-, R, R+
		3. Rain Showers	RW-, RW, RW+
		4. Drizzle or	L-, L, L+
		Freezing Drizzle	ZL-, ZL, ZL+
		5. Snow	S-, S, S+
		Snow Pellets	SP-, SP, SP+
		Snow Grains	SG-, SG, SG+
		Ice Crystals	IC-, IC, IC+
		6. Snow Showers	SW-, SW, SW+
		Snow Squalls	SQ-, SQ, SQ+
		7. Tornado	Tornado
		Thunderstorm	T, T+
		Squalls	Q-, Q, Q+
		Hail	A-, A, A+
		Hail Pellets	AP-, AP, AP+
		8. Freezing Rain	ZR-, ZR, ZR+
		Sleet	E-, E, E+
		Sleet Showers	EW-, EW, EW+
		9. Fog, Ice Fog	F, IF
		10. Ground Fog	GF
		11. Blowing Snow	BS
		Blowing Sand	BN
		Blowing Dust	BD
		12. Smoke, Haze, Dust	K, H, D
		Sky Condition, Lower:	
SCL	3	1. Sky obscured	X
		2. Thin broken clouds	- ⊕
		Thin overcast	- ⊕
		3. Scattered clouds	⊕
		Dark scattered clouds	+ ⊕

Table 1-5a

Selected Predictors and REEP Equations
Atlantic City 3-hr Ceiling

Order	Predictors Selected by MDA Station	Element	Dummy	Predictors Selected by REEP Station	Element	Dummy		The Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1	ACY	CIG	5	ACY	CIG	5		-.006	-.002	-.025	-.218	.251
2	PHL	CIG	5	ACY	CIG	1		.390	.192	-.025	-.359	-.199
3	ACY	CIG	4	PHL	CIG	5		.009	-.021	-.067	-.135	.213
4	ACY	CIG	3	ACY	CIG	2		.049	.318	-.002	-.289	-.076
5	DCA	CIG	5	ACY	CIG	3		-.000	.067	.188	-.211	-.043
6	ACY	CIG	1	DCA	CIG	5		.007	-.021	-.051	-.054	.119
7	SBY	WEA	2	SBY	CIG	5		-.004	-.061	-.036	-.008	.110
8	NEL	TCA	2	NEL	TCA	2		.005	.004	.014	.050	-.072
9	SBY	CIG	5	ACY	VIS	5		-.047	.011	-.010	-.003	.049
10	SBY	CIG	4	DCA	RLH	3		.013	.072	.001	-.052	-.034
11	SWF	WND	3	NEL	CIG	3		-.009	-.006	.090	-.029	-.046
12	DCA	RLH	3	ACY	CIG	(1)		-.177	.062	-.073	.048	.140
13	IDL	WND	2	SBY	CIG	4		-.010	-.078	-.026	.104	.010
14	ACY	VIS	5	IDL	WND	2		.010	.019	.019	.027	-.075
15	NEL	WEA	2	SBY	WEA	2		-.009	.078	-.007	.049	-.111
16	ACY	CIG	(1)	SWF	WND	3		.009	.021	.022	.018	-.071
17	IDL	CIG	3	SBY	VIS	2		.087	.011	-.037	-.036	-.026
18				ACY	CIG	(6)		-.010	-.014	-.039	.173	-.110
19				NEL	WEA	2		-.027	.015	.083	-.048	-.023
							Additive constant	.042	.089	.185	.429	.256

Table 1-5b

Selected Predictors and REEP Equations
Atlantic City 7-hr Ceiling

Order	Predictors Selected by MDA Station	Element	Dummy	Predictors Selected by REEP Station	Element	Dummy		The Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1	PHL	CIG	5	PHL	CIG	5		.002	-.032	-.051	-.087	.168
2	DCA	CIG	5	DCA	CIG	5		-.003	-.008	-.051	-.073	.135
3	ACY	CIG	5	ACY	CIG	5		-.015	-.030	-.038	-.061	.144
4	DCA	WEA	2	PHL	WND	3		.027	.023	.037	.045	-.132
5	IDL	WND	2	IDL	WND	2		.001	.033	.054	.039	-.127
6	PHL	WND	3	ACY	TCA	2		.003	.002	.027	.049	-.092
7	ACY	CIG	4	ACY	SCL	1		.104	-.043	-.041	-.033	.014
8	MDT	TCA	2	DCA	WEA	2		-.008	.057	.126	-.036	-.139
9	NEL	WEA	9	PSB	WEA	8		-.023	.253	-.090	-.042	-.099
10	ACY	TOD	1	NEL	WEA	9		.024	.016	.061	.009	-.110
11	MDT	WEA	2	ACY	CIG	2		.012	.093	.005	-.145	.035
12	PHL	WND	2	ACY	TOD	1		-.026	-.018	-.009	.025	.028
13	SBY	WEA	9	PSB	RLH	3		.007	-.007	.024	.041	-.064
14	ACY	SCL	1	DCA	CIG	3		-.001	.084	.009	-.049	-.043
15	PSB	WEA	8	SBY	WEA	9		.013	.047	.014	.031	-.104
16	RIC	WEA	2	PSB	CIG	1		.007	.092	-.037	-.040	-.021
17				PHL	WEA	7		-.031	.237	.029	-.063	-.172
							Additive constant	.031	.076	.137	.223	.533

Table 1-5c

Selected Predictors and REEP Equations
Atlantic City 3-hr Visibility

Order	Predictors Selected by MDA Station	Element	Dummy	Predictors Selected by REEP Station	Element	Dummy		The Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1	ACY	VIS	5	ACY	VIS	5		-.041	-.023	-.067	-.054	.185
2	ACY	VIS	1	ACY	VIS	1		.274	.038	-.017	-.156	-.139
3	ACY	RLH	3	ACY	RLH	3		.023	.018	.029	.021	-.090
4	ACY	VIS	2	ACY	VIS	2		.113	.178	-.010	-.076	-.204
5	DCA	WEA	9	DCA	WEA	9		.002	.012	.037	.049	-.101
6	SBY	RLH	3	SBY	VIS	5		-.004	-.021	-.030	-.060	.115
7	NEL	WEA	9	NEL	WEA	9		.017	.017	.023	.048	-.106
8	ACY	WEA	1	ACY	WEA	1		-.007	-.005	-.011	-.045	.068
9	PHL	CIG	1	PHL	CIG	1		.183	.018	-.079	-.084	-.039
10	NEL	WEA	2	NEL	WEA	2		-.057	-.015	.024	.033	.016
11	SBY	CIG	1	ORF	RLH	3		.040	.010	.005	.004	-.060
12	ORF	RLH	3	MDT	VIS	2		-.025	.020	.116	-.025	-.086
							Additive constant	.052	.051	.109	.165	.623

Table 1-5c

Selected Predictors and REEP Equations
Atlantic City 7-hr Visibility

Predictors Selected by MDA				Predictors Selected by REEP			The Coefficients of the REEP Equations				
Order	Station	Element	Dummy	Station	Element	Dummy	A_1	A_2	A_3	A_4	A_5
1	ACY	WEA	9	ACY	WEA	9	.052	.020	.050	.005	-.128
2	PHL	CIG	5	DCA	VIS	5	-.040	-.032	-.034	-.067	.173
3	ACY	TCD	1	ACY	TCD	1	-.037	-.019	-.026	-.027	.109
4	DCA	VIS	5	PHL	WEA	1	-.015	-.004	-.020	-.018	.055
5	DCA	WEA	2	PHL	CIG	5	.010	-.019	-.036	-.026	.071
6	NEL	WEA	1	IDL	VIS	3	.087	-.004	-.038	.012	-.058
7	ACY	VIS	5	PHL	WND	3	.018	.020	.021	.012	-.070
8	PHL	WND	5	ACY	VIS	2	.069	.099	-.015	-.037	-.115
9				DCA	WEA	2	-.022	-.016	.062	.076	-.101
10				NEL	WEA	1	-.011	-.009	-.010	-.026	.055
						Additive constant	.076	.075	.113	.156	.581

Table 1-6a

Contingency Tables Showing Predicted-Versus-Actual Groups for Persistence, MDA, and REEP

3-hr Ceiling (Independent Sample)

		Actual					
	Group	1	2	3	4	5	Total
PERSISTENCE	1	21	8	4	0	5	38
	2	13	33	25	10	9	90
	3	2	14	50	21	21	108
	4	1	8	21	68	70	168
	5	7	10	27	66	1530	1640
	Total	44	73	127	165	1635	2044
		Hits = 1702			$\bar{P}$ = 0.2385		
	$\bar{P}_g$	0.0156	0.0284	0.0462	0.0636	0.0847	

		Actual					
	Group	1	2	3	4	5	Total
MDA	1	13	5	2	0	4	24
	2	14	29	24	6	1	74
	3	1	13	41	15	6	76
	4	3	14	24	53	36	130
	5	13	12	36	91	1588	1740
	Total	44	73	127	165	1635	2044
		Hits = 1724			$\bar{P}$ = 0.2135		
	$\bar{P}_g$	0.0161	0.0279	0.0424	0.0601	0.0669	

		Actual					
	Group	1	2	3	4	5	Total
REEP	1	6	4	1	0	1	12
	2	19	34	25	8	4	90
	3	0	11	42	14	5	72
	4	1	6	19	41	20	87
	5	18	18	40	102	1605	1783
	Total	44	73	127	165	1635	2044
		Hits = 1728			$\bar{P}$ = 0.2135		
	$\bar{P}_g$	0.0163	0.0277	0.0423	0.0605	0.0666	

Table 1-6b

Contingency Tables Showing Predicted-Versus-Actual Groups for Persistence, MDA, and REEP

7-hr Ceiling (Independent Sample)

		Actual					
	Group	1	2	3	4	5	Total
PERSISTENCE	1	8	12	8	4	6	38
	2	12	22	19	19	18	90
	3	7	9	38	26	28	108
	4	2	12	33	37	84	168
	5	12	24	49	72	1483	1640
	Total	41	79	147	158	1619	2044
		Hits = 1588			$\bar{P}$ = 0.2899		
	$\bar{P}_g$	0.0183	0.0342	0.0598	0.0671	0.1105	

		Actual					
	Group	1	2	3	4	5	Total
MDA	1	0	0	0	0	0	0
	2	6	9	10	4	0	29
	3	9	29	42	15	12	107
	4	2	5	9	13	12	41
	5	24	36	86	126	1595	1867
	Total	41	79	147	158	1619	2044
		Hits = 1659			$\bar{P}$ = 0.2589		
	$\bar{P}_g$	0.0181	0.0321	0.0539	0.0653	0.0894	

		Actual					
	Group	1	2	3	4	5	Total
REEP	1	0	0	0	0	0	0
	2	3	9	9	1	2	24
	3	7	19	34	12	9	81
	4	8	7	22	20	6	63
	5	23	44	82	125	1602	1876
	Total	41	79	147	158	1619	2044
		Hits = 1665			$\bar{P}$ = 0.2600		
	$\bar{P}_g$	0.0180	0.0323	0.0540	0.0666	0.0891	

Table 1-6c

Contingency Tables Showing Predicted-Versus-Actual Groups for Persistence, MDA, and REEP

3-hr Visibility (Independent Sample)

		Actual					
	Group	1	2	3	4	5	Total
PERSISTENCE	1	17	10	7	1	6	41
	2	4	8	9	1	16	38
	3	4	11	21	11	31	78
	4	3	6	13	23	45	90
	5	14	11	28	57	1687	1797
	Total	42	46	78	93	1785	2044
		Hits = 1756		$\bar{P}$ = 0.1910			
	$\bar{P}_g$	0.0170	0.0200	0.0334	0.0413	0.0793	

		Actual					
	Group	1	2	3	4	5	Total
MDA	1	16	10	6	1	5	38
	2	2	4	3	0	5	14
	3	0	0	0	0	0	0
	4	0	0	0	0	0	0
	5	24	32	69	92	1775	1992
	Total	42	46	78	93	1785	2044
		Hits = 1795		$\bar{P}$ = 0.1795			
	$\bar{P}_g$	0.0173	0.0201	0.0329	0.0391	0.0701	

		Actual					
	Group	1	2	3	4	5	Total
REEP	1	13	9	5	1	4	32
	2	2	5	4	0	9	20
	3	0	2	0	0	0	2
	4	0	0	3	4	0	7
	5	27	30	66	88	1772	1983
	Total	42	46	78	93	1785	2044
		Hits = 1794		$\bar{P}$ = 0.1800			
	$\bar{P}_g$	0.0169	0.0199	0.0331	0.0395	0.0707	

Table 1-6d

Contingency Tables Showing Predictand-Versus-Actual Groups for Persistence, MDA, and REEP

7-hr Visibility (Independent Sample)

		Actual					
	Group	1	2	3	4	5	Total
PERSISTENCE	1	6	4	6	7	18	41
	2	3	3	4	4	24	38
	3	4	8	17	9	40	78
	4	7	2	15	11	55	90
	5	23	21	48	68	1637	1797
	Total	43	38	90	99	1774	2044
		Hits = 1674		$\bar{P}$ = 0.2270			
	$\bar{P}_g$	0.0201	0.0180	0.0405	0.0455	0.1029	

		Actual					
	Group	1	2	3	4	5	Total
MDA	1	0	0	0	0	0	0
	2	0	0	0	0	0	0
	3	0	0	0	0	0	0
	4	0	0	0	0	0	0
	5	43	38	90	99	1774	2044
	Total	43	38	90	99	1774	2044
		Hits = 1774		$\bar{P}$ = 0.2120			
	$\bar{P}_g$	0.0191	0.0178	0.0391	0.0448	0.0912	

		Actual					
	Group	1	2	3	4	5	Total
REEP	1	0	0	1	0	1	2
	2	0	0	0	0	0	0
	3	0	0	0	0	0	0
	4	0	0	0	0	0	0
	5	43	38	89	99	1773	2042
	Total	43	38	90	99	1774	2044
		Hits = 1773		$\bar{P}$ = 0.2150			
	$\bar{P}_g$	0.0196	0.0178	0.0394	0.0446	0.0936	

Table 1-7

Comparison Between Persistence, MDA, and REEP

(Independent Sample - 2044 Observations)

Atlantic City, Ceiling 3 Hours in Advance

Hits	Persistence	1702
	MDA	1724
	REEP	1728*
$\bar{P}$ Score	Persistence	.2385
	MDA	.2135*
	REEP	.2135*

Atlantic City, Ceiling 7 Hours in Advance

Hits	Persistence	1588
	MDA	1659
	REEP	1665*
$\bar{P}$ Score	Persistence	.2899
	MDA	.2589*
	REEP	.2600

Atlantic City, Visibility 3 Hours in Advance

Hits	Persistence	1756
	MDA	1795*
	REEP	1794
$\bar{P}$ Score	Persistence	.1910
	MDA	.1795*
	REEP	.1800

Atlantic City, Visibility 7 Hours in Advance

Hits	Persistence	1674
	MDA	1774*
	REEP	1773
$\bar{P}$ Score	Persistence	.2270
	MDA	.2120*
	REEP	.2150

Asterisks are used to denote superiority.

EXPERIMENT 1A. Comparison Between MDA and REEP (Offutt AFB, Nebraska)

The Purpose and the Procedure are the same as those in Experiment No. 1.

Predictands

Visibility at Offutt AFB was divided into four and ceiling into five operationally significant categories (see Table 1A-1) for the following four predictands: CIG + 3 hours, CIG + 7 hours, VIS + 3 hours, and VIS + 7 hours. The specifications of the predictands are tabulated in Table 1A-2.

Predictors

As in Experiment 1, seven meteorological elements at the 13 predictor stations constituted the Offutt AFB network (see Figure 1A-1), plus the time of day and day of year variables. These were then transformed into 423 dummy variables (see Tables 1A-3 and 1A-4).

Data Sample

A data sample containing 10,000 observations was selected at random from the ten year period, 1 January 1949 to 31 December 1958. The developmental sample from the first eight years contained 7668 observations, and the verification sample from the last two years contained 2332 observations.

Results

The four sets of predictors selected by MDA and REEP are shown in Tables 1A-5a to 1A-5d in addition to the REEP regression coefficients for each group of the predictand. The contingency tables (see Tables 1A-6a to 1A-6d) give the forecast-versus-observed frequencies for

persistence, MDA, and REEP. The number of hits, the $\bar{P}$ scores, and the individual group contributions to the $\bar{P}$ score are also shown. Table 1A-7 summarizes the number of correct forecasts and the overall $\bar{P}$ scores.

Conclusions

1) In the Atlantic City experiment (Experiment 1) the selected predictors for REEP and MDA are both alike in numbers and in kind. However, in the Offutt experiment (Experiment 1A) this is not the case. There are many more predictors selected as significant in REEP. This is due to two adjustable factors introduced into the REEP computer program for the Offutt experiment but not for the Atlantic experiment.

The first of these is that in REEP each group can be assigned its own effective degrees of freedom determinable from a runs test, whereas MDA must use an average number (from previous studies determined empirically to be at about one-third the developmental sample size or roughly crediting only one observation a day as independent). Most of these additional selected predictors explain a significant amount of variability in the low ceiling and visibility categories.

Another adjustable factor is that of minimum frequency in the number of observations for a possible predictor (in our terminology "bit count"). In order for a predictor to be selected it must have been observed at least a minimum number of times in the developmental sample. This is an adjustable input parameter to the

program*. A more detailed discussion on its effects is given in a subsequent experiment. Suffice it to say that for larger values it selects more stable predictors with a tendency toward predictor selection to favor higher ceiling and visibility categories. An additional effect is that for larger values it would select fewer predictors thus offsetting some of the effects due to the degrees of freedom factor.

2) A summary of the results shows MDA with ten asterisks in its favor while REEP has eight. In almost every instance where an individual $\overline{P}$ score or the number of hits differ this difference is very small. We conclude that for all practical purposes MDA and REEP produce comparable results.

*The present MDA program now has this as an input parameter. At the time of the Atlantic City and Offutt comparisons it was not available.

Table 1A-1

Definition of Predictand Groups for Offutt AFB, Nebraska

Group Number	Ceiling Intervals
1	0 ft $\leq$ Ceiling $<$ 300 ft
2	300 ft $\leq$ Ceiling $<$ 1000 ft
3	1000 ft $\leq$ Ceiling $<$ 1500 ft
4	1500 ft $\leq$ Ceiling $<$ 5000 ft
5	5000 ft $\leq$ Ceiling

Group Number	Visibility Intervals
1	0 mi $\leq$ Visibility $<$ 1 mi
2	1 mi $\leq$ Visibility $<$ 3 mi
3	3 mi $\leq$ Visibility $<$ 5 mi
4	5 mi $\leq$ Visibility

Table 1A-2

Specification of Predictands - Offutt AFB, Nebraska

		Predictand 1	2	3	4
Predictand variable		CIG	CIG	VIS	VIS
Forecast interval--hours (H)		3	7	3	7
Number of Groups (G)		5	5	4	4
Observations in Group 1 (n_1)		49	63	94	106
Observations in Group 2 (n_2)		373	379	212	213
Observations in Group 3 (n_3)		236	259	250	223
Observations in Group 4 (n_4)		959	920	7112	7126
Observations in Group 5 (n_5)		6051	6047	-	-
Total dependent sample size (N)		7668	7668	7668	7668
Number of independent observations in the dependent sample (n)		2256	2256	2256	2256
Degrees of Freedom for each group n_g ($g=1,\ldots,G$)	Group 1	5400	5400	5500	5500
	Group 2	4600	4600	5800	5800
	Group 3	5800	5800	6300	6300
	Group 4	5200	5200	4400	4400
	Group 5	3500	3500	-	-
Total independent sample size (M)		2332	2332	2332	2332
Number of available predictors (P)		423	423	423	423

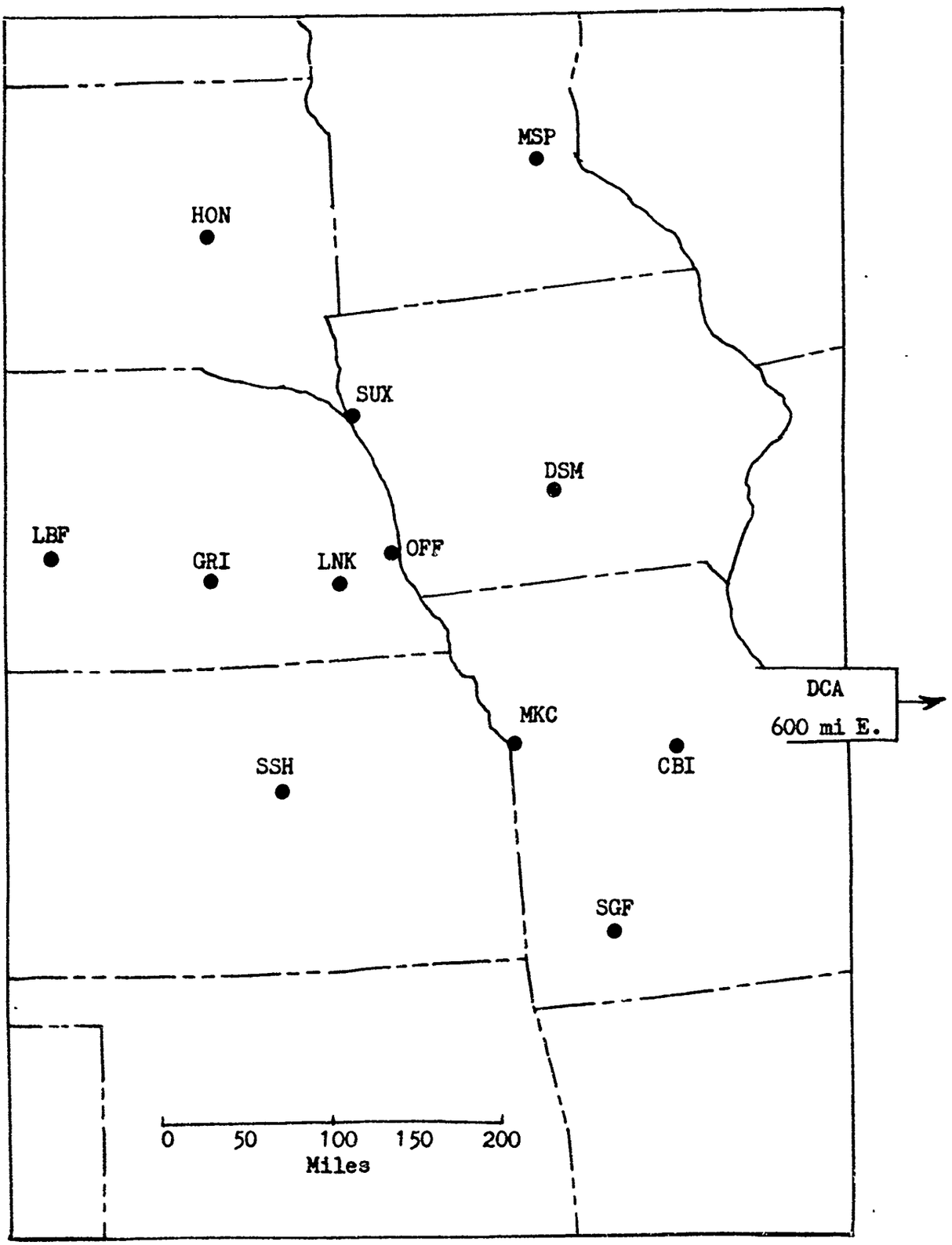

Figure 1A-1. Thirteen Station Network for Offutt AFB, Nebraska

Table 1A-3

Number of Possible Predictors for the
3- and 7-hr Ceiling and Visibility at OFF

Elem.	Station OFF	GRI	SUX	DSM	MKC	SSH	LBF	HON	MSP	LNK	CBI	SGF	DCA	Total
TOD	5	-	-	-	-	-	-	-	-	-	-	-	-	5
DOY	1	-	-	-	-	-	-	-	-	-	-	-	-	1
CIG	15	5	5	5	5	5	5	5	5	5	5	5	-	70
VIS	14	4	4	4	4	4	4	4	4	4	4	4	-	58
WND	5	5	5	5	5	5	5	5	5	5	5	5	5	65
RLH	3	3	3	3	3	3	3	3	3	3	3	3	3	39
TCA	2	2	2	2	2	2	2	2	2	2	2	2	2	26
WEA	12	12	12	12	12	12	12	12	12	12	12	12	12	156
SCL	3	-	-	-	-	-	-	-	-	-	-	-	-	3
													Total	423

Table 1A-4

Specifications and Limits of the Zero-One (Dummy) Variables for Offutt AFB, Nebraska

Meteorological Element	Number of Dummies	Limits of Each Category
		Time of Day:
TOD	5	(Same as for ACY - see Table 1-4)
		Day of Year:
DOY	1	(Same as for ACY - see Table 1-4)
		Wind:
WND	5	(Same as for ACY - see Table 1-4)
		Relative Humidity:
RLH	3	(Same as for ACY - see Table 1-4)
		Total Cloud Amount:
TCA	2	(Same as for ACY - see Table 1-4)
		Ceiling Height, ft:
CIG	5	1. $0 \leq CIG < 300$
		2. $300 \leq CIG < 1000$
		3. $1000 \leq CIG < 1500$
		4. $1500 \leq CIG < 5000$
		5. $5000 \leq CIG$
		Ceiling Height*, ft:
CIG	10	(1) $CIG = 100$
		(2) $CIG = 200$
		(3) $CIG = 400$
		(4) $CIG = 500$
		(5) $CIG = 900$

*Parentheses are used here to help distinguish the two forms of ceiling and visibility dummies.

Table 1A-4 - Continued

Meteorological Element	Number of Dummies	Limits of Each Category
		Ceiling Height*, ft:
		(6) $1500 \leq CIG < 2000$
		(7) $2500 \leq CIG < 3000$
		(8) $3000 \leq CIG < 5000$
		(9) $5000 \leq CIG < 10{,}000$
		(10) CIG = Unlimited
		Visibility, mi:
VIS	4	1. $0 \leq VIS < 1$
		2. $1 \leq VIS < 3$
		3. $3 \leq VIS < 5$
		4. $5 \leq VIS$
		Visibility*, mi:
VIS	10	(1) $1/4 \leq VIS < 1/2$
		(2) $1/2 \leq VIS < 3/4$
		(3) $3/4 \leq VIS < 1$
		(4) $1 \leq VIS < 1\text{-}1/2$
		(5) $1\text{-}1/2 \leq VIS < 2$
		(6) $2 \leq VIS < 3$
		(7) $3 \leq VIS < 4$
		(8) $4 \leq VIS < 5$
		(9) $5 \leq VIS < 6$
		(10) $15 \leq VIS$
		Weather Element:
WEA	12	(Same as for ACY - see Table 1-4)
		Sky Condition, Lower:
SCL	3	(Same as for ACY - see Table 1-4)

*Parentheses are used here to help distinguish the two forms of ceiling and visibility dummies.

Table 1A-5a

Selected Predictors and REEP Equations
Offutt AFB, Nebraska 3-hr Ceiling

Order	Predictors Selected by MDA Station	Element	Dummy	Predictors Selected by REEP Station	Element	Dummy	The Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1	OFF	CIG	5	OFF	CIG	5	-.413	-.123	.045	.135	.356
2	OFF	CIG	4	OFF	CIG	2	-.388	.192	.090	.060	.047
3	OFF	CIG	1	LNK	CIG	5	-.260	.145	.055	-.092	.153
4	SUX	CIG	5	OFF	VIS	1	-.018	-.090	.059	.106	-.057
5	LNK	CIG	5	SUX	CIG	5	-.003	-.118	-.057	-.054	.231
6	OFF	CIG	2	LNK	CIG	2	-.270	.260	.095	-.077	-.008
7	SUX	CIG	4	OFF	WEA	9	.021	.024	.009	-.018	-.036
8	GRI	CIG	5	LNK	CIG	3	-.273	.105	.237	-.005	-.065
9	LNK	CIG	1	LNK	CIG	4	-.262	.144	.088	.044	-.014
10	LNK	CIG	2	OFF	CIG	3	-.421	-.065	.198	.260	.028
11	DSM	WEA	4	OFF	CIG	4	-.416	-.141	.046	.455	.056
12	GRI	CIG	4	GRI	CIG	5	-.006	-.056	-.025	-.056	.143
13	LNK	TCA	2	SUX	CIG	4	.002	-.103	-.037	.100	.038
14	LNK	RLH	3	DSM	VIS	2	.030	.009	-.049	.043	-.033
15	LNK	CIG	4	OFF	WEA	5	-.008	-.098	.121	.004	-.019
16	DSM	WEA	9	LNK	VIS	4	.008	-.091	.073	.057	-.047
17	OFF	VIS	(7)	LNK	TCA	2	.000	.015	.016	.041	-.072
18	SSH	CIG	1	DSM	CIG	4	-.001	-.004	-.006	.065	-.055
19	DSM	CIG	5	GRI	CIG	4	-.006	-.068	-.027	.065	.035
20				LNK	WEA	2	-.016	.094	.019	-.018	-.079
21				SGF	CIG	1	-.010	.091	-.049	-.065	.033
22				DSM	CIG	2	-.011	.052	.018	-.013	-.045
23				DSM	WEA	9	.029	.004	.009	-.026	-.016
24				GRI	VIS	1	.038	-.002	.057	-.089	-.004
25				HON	WEA	9	.019	-.032	.064	-.040	-.011
26				OFF	CIG	(6)	-.000	.026	.065	-.057	-.033
27				LBF	CIG	5	.003	-.005	.015	-.056	.043
28				DSM	WEA	2	-.030	.046	-.011	.008	-.013
29				OFF	CIG	(8)	.003	.012	-.001	-.102	.087
30				SSH	WEA	9	.004	.061	-.018	-.027	-.020
					Additive Constant:		.671	.247	-.107	.082	.107

Table 1A-5b

Selected Predictors and REEP Equations
Offutt AFB, Nebraska 7-hr Ceiling

Predictors Selected by MDA				Predictors Selected by REEP			The Coefficients of the REEP Equations				
Order	Station	Element	Dummy	Station	Element	Dummy	A_1	A_2	A_3	A_4	A_5
1	LNK	CIG	5	LNK	CIG	5	.013	-.021	-.036	-.045	.090
2	SUX	CIG	5	SUX	CIG	5	-.021	-.123	-.037	.030	.152
3	LNK	WEA	9	OFF	CIG	2	-.025	.090	.034	-.065	-.034
4	GRI	CIG	5	GRI	CIG	5	-.136	.097	-.078	-.103	.219
5	SUX	CIG	4	DSM	WEA	9	.023	.024	-.016	-.007	-.025
6	OFF	CIG	1	GRI	CIG	2	-.135	.199	-.034	-.037	.006
7	HON	TCA	2	OFF	SCL	1	.081	-.022	-.045	-.013	-.001
8	DSM	RLH	3	HON	TCA	2	-.001	.006	.012	.047	-.063
9	GRI	CIG	4	LNK	WEA	9	.050	.079	-.063	-.090	.025
10	OFF	TCA	2	SSH	CIG	2	-.020	.106	-.014	-.014	-.058
11	GRI	CIG	1	SUX	CIG	4	-.019	-.100	-.013	.110	.021
12	SSH	CIG	2	DSM	WEA	5	-.015	-.048	-.001	.165	-.102
13	CBI	WEA	9	CBI	WEA	9	.039	-.019	.029	-.021	-.028
14	OFF	CIG	2	OFF	TCA	2	.004	.004	.017	.070	-.095
15	HON	CIG	5	GRI	VIS	1	.012	.092	.008	-.113	.001
16	OFF	TOD	3	MSP	CIG	2	.039	-.002	.006	-.000	-.043
17	DSM	WEA	5	GRI	CIG	4	-.138	.081	-.079	.058	.077
18				GRI	CIG	3	-.142	.096	.007	-.009	.048
19				HON	CIG	5	-.003	-.006	-.009	-.074	.092
20				MKC	CIG	4	-.001	-.013	-.003	.072	-.055
21				OFF	TOD	3	-.009	-.032	-.010	.020	.031
22				DSM	VIS	2	.034	-.018	-.036	.023	-.003
23				LNK	WEA	2	.002	.112	-.029	-.016	-.069
24				SUX	CIG	3	-.014	-.010	.077	.047	-.100
25				LBF	CIG	5	.001	-.014	-.010	-.059	.082
26				MKC	WEA	9	.035	.003	.015	-.037	-.015
27				DSM	WND	3	.001	.029	-.002	.007	-.034
28				OFF	CIG	3	-.024	.009	.062	-.007	-.040

Table 1A-5b - Continued

Predictors Selected by MDA				Predictors Selected by REEP			The Coefficients of the REEP Equations				
Order	Station	Element	Dummy	Station	Element	Dummy	A_1	A_2	A_3	A_4	A_5
29				CBI	WEA	1	.003	-.041	-.000	.013	.024
30				LNK	VIS	3	-.040	-.005	.058	.050	-.063
31				OFF	VIS	3	.067	.014	.003	.012	-.097
32				OFF	VIS	(8)	-.062	.003	.004	.003	.053
33				MSP	WEA	1	.004	-.004	.006	-.050	.044
34				GRI	VIS	2	.002	.095	-.012	-.079	-.005
35				DSM	RLH	3	.009	.055	.022	-.014	-.072
36				CBI	CIG	1	-.044	.002	-.008	.003	.047
						Additive Constant:	.141	.111	.166	.295	.288

Table 1A-5c

Selected Predictors and REEP Equations
Offutt AFB, Nebraska 3-hr Visibility

Order	Predictors Selected by MDA Station	Element	Dummy	Predictors Selected by REEP Station	Element	Dummy	The Coefficients of the REEP Equations A_1	A_2	A_3	A_4
1	OFF	VIS	4	OFF	VIS	4	.005	-.096	-.228	.319
2	OFF	CIG	1	OFF	VIS	1	.235	.014	-.178	-.071
3	GRI	VIS	4	LNK	WEA	1	-.011	-.016	-.035	.062
4	SUX	VIS	4	SUX	VIS	4	-.004	-.075	-.041	.120
5	OFF	VIS	3	GRI	VIS	4	-.015	-.048	-.011	.074
6	SSH	CIG	1	OFF	VIS	2	.045	.149	-.163	-.032
7	LNK	WEA	1	OFF	VIS	(9)	-.013	.019	.126	-.132
8	LNK	VIS	1	GRI	VIS	1	.093	-.069	.034	-.058
9	OFF	VIS	1	MKC	WEA	9	.016	.039	.025	-.080
10	SSH	WEA	8	OFF	SCL	1	.119	-.135	-.003	.018
11	OFF	VIS	(9)	LNK	VIS	2	-.023	.098	-.019	-.056
12	CBI	WEA	9	OFF	WEA	9	.056	-.007	-.033	-.016
13	OFF	VIS	(2)	OFF	CIG	2	-.054	.032	-.003	.025
14	DSM	CIG	1	SUX	WEA	9	.052	.001	.015	-.067
15	GRI	WEA	5	GRI	WEA	5	.029	.067	.042	-.138
16	OFF	VIS	(3)	SUX	WEA	2	-.060	-.003	.002	.060
17	OFF	VIS	(1)	CBI	WEA	2	.043	.021	-.000	-.064
18	LNK	WEA	9	LNK	CIG	2	-.020	.014	.076	-.070
19	GRI	VIS	1	OFF	WEA	5	-.046	-.014	-.007	.067
20				SSH	RLH	3	.031	.029	-.013	-.048
21				OFF	TOD	3	-.009	-.009	-.011	.029
22				MSP	WEA	9	.009	.018	.031	-.059
23				LNK	WEA	2	.017	-.052	.081	-.046
24				LNK	RLH	3	-.006	.045	.001	-.040
25				DSM	VIS	2	.042	-.048	-.013	.019
26				LBF	VIS	3	-.007	.067	-.028	-.032
27				DSM	WEA	5	-.035	.025	.021	-.011
28				SGF	CIG	1	.021	.064	-.057	-.028
29				OFF	RLH	2	.002	.005	.027	-.034
					Additive Constant:		.026	.235	.322	.417

Table 1A-5d

Selected Predictors and REEP Equations
Offutt AFB, Nebraska 7-hr Visibility

Order	Predictors Selected by MDA Station	Element	Dummy	Predictors Selected by REEP Station	Element	Dummy	The Coefficient of the REEP Equations A_1	A_2	A_3	A_4
1	OFF	VIS	4	OFF	VIS	4	-.036	-.034	-.044	.114
2	GRI	VIS	4	GRI	VIS	4	-.070	-.083	-.009	.163
3	SUX	WEA	1	SUX	WEA	1	-.001	-.047	-.034	.082
4	DSM	WEA	9	DSM	WEA	9	.052	.012	.018	-.081
5	LNK	VIS	1	OFF	SCL	1	.086	-.025	-.067	.006
6	LBF	WEA	5	LBF	WEA	5	.042	.068	.002	-.113
7	SSH	CIG	1	SUX	WEA	9	.010	.082	.028	-.121
8	SUX	WEA	9	GRI	VIS	3	-.081	-.019	.014	.085
9	SSH	RLH	2	MKC	WEA	9	.041	-.031	.015	-.025
10				LNK	TCA	2	-.007	-.001	.026	-.019
11				MSP	CIG	2	.044	-.010	-.013	-.021
12				SGF	VIS	4	.009	-.040	-.007	.038
13				OFF	TOD	3	-.009	-.016	-.018	.043
14				LNK	CIG	2	-.037	.066	.022	-.051
15				SSH	RLH	3	.040	.050	-.035	-.056
16				SSH	RLH	2	.004	.029	.024	-.057
17				DSM	WEA	2	-.048	-.029	.001	.076
18				SUX	VIS	2	.036	-.016	.073	-.093
19				SUX	RLH	3	.024	-.050	-.013	.038
20				DSM	WEA	12	.016	.033	.040	-.089
21				LNK	WEA	2	-.040	.001	.007	-.039
22				LBF	CIG	2	-.018	.043	.017	-.042
23				LNK	WND	5	-.013	-.011	-.005	.030
24				CBI	VIS	2	.039	.010	-.021	-.028
				Additive Constant:			.104	.212	.106	.577

Table 1A-6a

Contingency Tables Showing Predicted-Versus-Actual Groups for Persistence, MDA, and REEP

OFF 3-hr Ceiling (Independent Sample)

PERSISTENCE

	Actual					
Group	1	2	3	4	5	Total
1	9	6	0	0	3	18
2	4	64	22	19	8	117
3	3	20	18	21	8	70
4	0	21	16	199	119	355
5	1	11	16	122	1622	1772
Total	17	122	72	361	1760	2332

Hits = 1912 $\quad \bar{P} = .2664$

$\bar{P}_g$	.0053	.0338	.0270	.1000	.1003

MDA

	Actual					
Group	1	2	3	4	5	Total
1	9	7	0	0	3	19
2	5	75	32	21	7	140
3	0	7	7	7	2	23
4	0	22	15	183	70	290
5	3	11	18	150	1678	1860
Total	17	122	72	361	1760	2332

Hits = 1952 $\quad \bar{P} = .2369$

$\bar{P}_g$	.0054	.0310	.0271	.0888	.0845

REEP

	Actual					
Group	1	2	3	4	5	Total
1	6	5	0	0	2	13
2	5	74	30	22	5	136
3	0	6	5	3	1	15
4	0	23	16	177	62	278
5	6	14	21	159	1690	1890
Total	17	122	72	361	1760	2332

Hits = 1952 $\quad \bar{P} = .2368$

$\bar{P}_g$	.0057	.0304	.0265	.0897	.0845

Table 1A-6b

Contingency Tables Showing Predicted-Versus-Actual Groups for Persistence, MDA, and REEP

OFF 7-hr Ceiling (Independent Sample)

		Actual					
	Group	1	2	3	4	5	Total
PERSISTENCE	1	4	7	0	2	5	18
	2	3	45	20	28	21	117
	3	3	16	10	20	21	70
	4	0	26	21	134	174	355
	5	6	23	25	170	1548	1772
	Total	16	117	76	354	1769	2332
		Hits = 1741			$\bar{P}$ = .3319		
	$\bar{P}_g$	.0065	.0392	.0301	.1185	.1376	

		Actual					
	Group	1	2	3	4	5	Total
MDA	1	1	2	0	0	1	4
	2	7	57	28	41	20	153
	3	1	0	1	1	1	4
	4	0	27	17	94	53	191
	5	7	31	30	218	1694	1980
	Total	16	117	76	354	1769	2332
		Hits = 1847			$\bar{P}$ - .2906		
	$\bar{P}_g$	.0065	.0369	.0298	.1090	.1084	

		Actual					
	Group	1	2	3	4	5	Total
REEP	1	1	0	0	1	2	4
	2	6	50	19	28	11	114
	3	0	2	1	0	0	3
	4	1	29	20	82	44	176
	5	8	36	36	243	1712	2035
	Total	16	117	76	354	1769	2332
		Hits = 1846			$\bar{P}$ = .2890		
	$\bar{P}_g$	.0069	.0352	.0285	.1095	.1089	

Table 1A-6c

Contingency Tables Showing Predicted-Versus-Actual Groups for Persistence, MDA, and REEP

OFF 3-hr Visibility (Independent Sample)

PERSISTENCE

	Actual				
Group	1	2	3	4	Total
1	24	11	7	6	48
2	12	18	16	17	63
3	4	14	18	32	68
4	6	20	39	2088	2153
Total	46	63	80	2143	2332

Hits = 2148 $\quad \bar{P} = .1093$

$\bar{P}_g$	.0140	.0227	.0300	.0425

MDA

	Actual				
Group	1	2	3	4	Total
1	13	5	3	3	24
2	15	23	14	8	60
3	3	7	6	8	24
4	15	28	57	2124	2224
Total	46	63	80	2143	2332

Hits = 2166 $\quad \bar{P} = .0996$

$\bar{P}_g$	.0138	.0214	.0286	.0358

REEP

	Actual				
Group	1	2	3	4	Total
1	17	9	5	3	34
2	9	17	11	4	41
3	2	5	7	6	20
4	18	32	57	2130	2237
Total	46	63	80	2143	2332

Hits = 2171 $\quad \bar{P} = .1007$

$\bar{P}_g$	.0141	.0216	.0287	.0363

Table 1A-6d

Contingency Tables Showing Predicted-Versus-Actual Groups for Persistence, MDA, and REEP

OFF 7-hr Visibility (Independent Sample)

		Actual				
	Group	1	2	3	4	Total
PERSISTENCE	1	14	8	5	21	48
	2	8	8	13	34	63
	3	6	9	17	36	68
	4	20	40	60	2033	2153
	Total	48	65	95	2124	2332
		Hits = 2072			$\bar{P}$ = .1504	
	$\bar{P}_g$	.0183	.0261	.0374	.0686	

		Actual				
	Group	1	2	3	4	Total
MDA	1	5	3	4	3	15
	2	6	5	9	9	29
	3	0	0	0	0	0
	4	37	57	82	2112	2288
	Total	48	65	95	2124	2332
		Hits = 2122			$\bar{P}$ = .1411	
	$\bar{P}_g$	.0184	.0256	.0372	.0600	

	Group	1	2	3	4	Total
REEP	1	2	1	1	5	9
	2	7	4	5	4	20
	3	0	0	0	0	0
	4	39	60	89	2115	2303
	Total	48	65	95	2124	2332
		Hits = 2121			$\bar{P}$ = .1402	
	$\bar{P}_g$	.0181	.0255	.0368	.0598	

Table 1A-7

Comparison Between Persistence, MDA, and REEP

(Independent Sample - 2332 Observations)

Offutt AFB, Ceiling 3 Hours in Advance

Hits	Persistence	1912
	MDA	1952*
	REEP	1952*
$\bar{P}$ Score	Persistence	.2664
	MDA	.2369
	REEP	.2368*

Offutt AFB, Ceiling 7 Hours in Advance

Hits	Persistence	1741
	MDA	1847*
	REEP	1846
$\bar{P}$ Score	Persistence	.3319
	MDA	.2906
	REEP	.2890*

Offutt AFB, Visibility 3 Hours in Advance

Hits	Persistence	2148
	MDA	2166
	REEP	2171*
$\bar{P}$ Score	Persistence	.1093
	MDA	.0996*
	REEP	.1007

Offutt AFB, Visibility 7 Hours in Advance

Hits	Persistence	2072
	MDA	2122*
	REEP	2121
$\bar{P}$ Score	Persistence	.1504
	MDA	.1411
	REEP	.1402*

Asterisks are used to denote superiority.

Purpose

A test to compare the results obtained when the predictor variables are in raw dummy form to those obtained when the predictor dummies are in cumulative form.

Predictands

Predictands used in this experiment were the Atlantic City ceilings for periods of 2, 5, and 7 hours in advance. These three predictands were divided into five intervals according to the operationally significant categories of ceiling, as shown in Table 2-1. Specifications for each predictand are shown in Table 2-2.

Predictors

Seven meteorological elements at ten predictor stations (see Figure 2-1) plus the time of day (TOD) and day of year (DOY) variables were transformed into dummy variables by dividing each predictor element into a number of intervals or categories (Table 2-3).

Data Sample

The data used for this experiment is the same as was used for Experiment No. 1, consisting of 10,000 observations selected randomly from the ten-year period from 1 January 1949 to 31 December 1958. The developmental sample consists of 7956 observations from the first eight years and the verification or independent sample contains 2044 observations from the last two years.

The set of 450 standard form dummy predictors are shown in Table 2-3. Table 2-4 tabulates the cumulative dummy predictors. Table 2-5 gives the specifications and limits of the dummy variables when used in the standard form and Table 2-6 gives similar information when the predictors are in the cumulative dummy form. Predictions using REEP were made for Atlantic City ceiling at 2, 5, and 7 hours in advance using the standard dummies and the cumulative dummies separately.

Results

The dummy predictors selected by REEP are shown in Tables 2-7a to 2-7f, as well as the REEP regression equation coefficients for each group of the predictand. Results of the forecasts using the two sets of predictors are shown in the contingency tables (Tables 2-8a to 2-8c) for the independent sample. Also shown are the results using persistence. The number of correct forecasts (hits), the Brier and Allen $\bar{P}$ scores, and the individual group contributions to the overall $\bar{P}$ score are given below each contingency table. The number of hits and overall $\bar{P}$ scores are summarized in Table 2-9.

Conclusions

From these limited results there does not appear to be any reason to prefer one form of dummy variables over the other.

Definition of Predictand Groups for Atlantic City Municipal Airport

Group Number	
1	0 ft $\leq$ Ceiling $<$ 200 ft
2	200 ft $\leq$ Ceiling $<$ 500 ft
3	500 ft $\leq$ Ceiling $<$ 1000 ft
4	1000 ft $\leq$ Ceiling $<$ 3000 ft
5	3000 ft $\leq$ Ceiling

Table 2-2

Specification of Predictands
Atlantic City Municipal Airport

	Predictand 1	2	3
Predictand variable	CIG	CIG	CIG
Forecast interval--hours (H)	2	5	7
Number of Groups (G)	5	5	5
Observations in Group 1 (n_1)	116	128	135
Observations in Group 2 (n_2)	263	259	272
Observations in Group 3 (n_3)	395	426	420
Observations in Group 4 (n_4)	656	631	619
Observations in Group 5 (n_5)	6526	6512	6510
Total dependent sample size (N)	7956	7956	7956
Number of independent observations in the dependent sample (n)	2652	2652	2652
Total independent sample size (M)	2044	2044	2044
Number of available predictors (P)	378	378	378

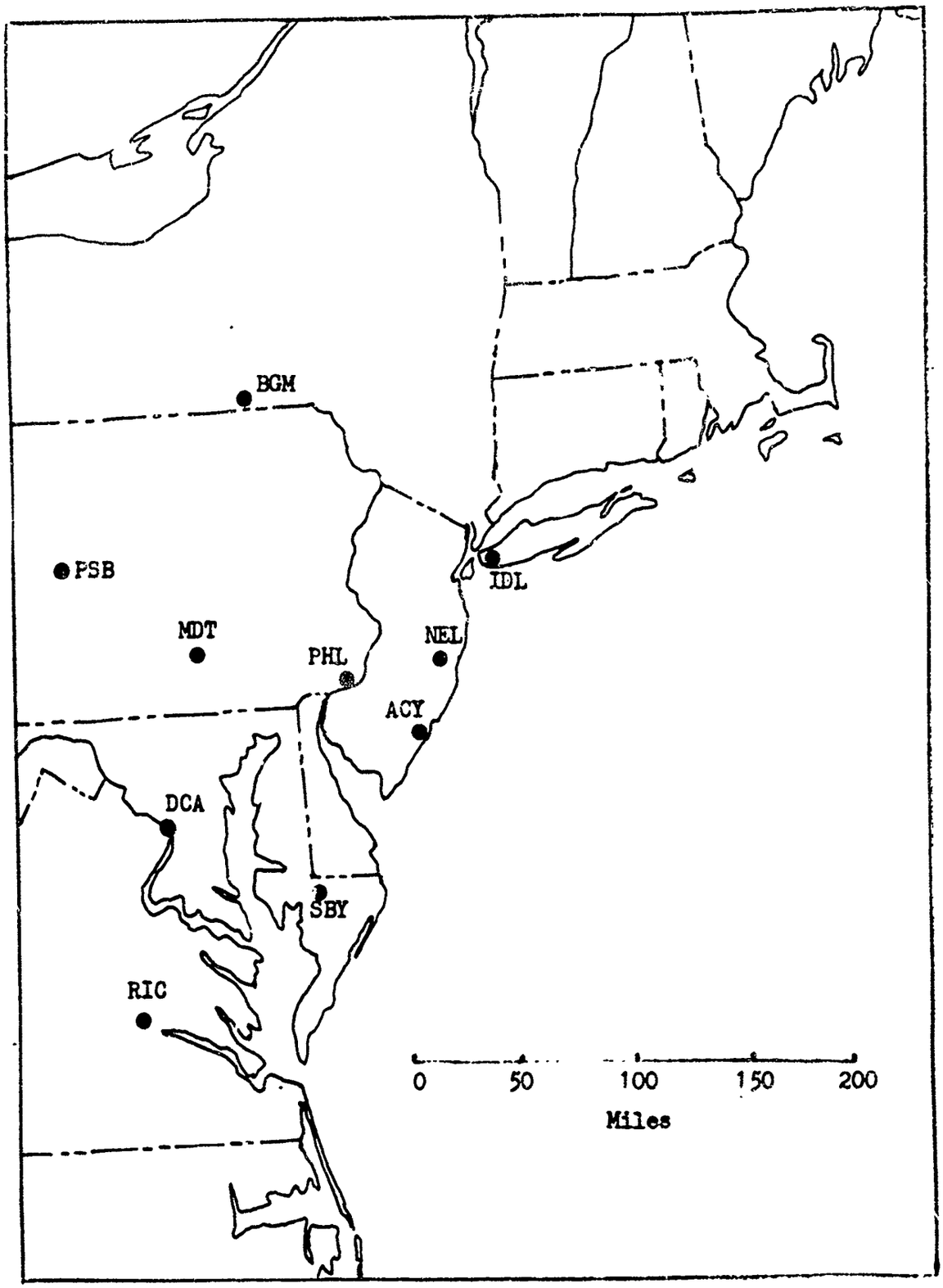

Figure 2-1. Ten Station Network for Atlantic City Municipal Airport

Table 2-3

Number of Possible Predictors When Used in Standard Form

Elem.	Station ACY	NEL	PHL	SBY	MDT	DCA	PSB	BGM	RIC	IDL	Total
TOD	1	-	-	-	-	-	-	-	-	-	1
DOY	1	-	-	-	-	-	-	-	-	-	1
CIG	24	5	5	5	5	5	5	5	5	5	69
VIS	24	5	5	5	5	5	5	5	5	5	69
UWC	5	5	5	5	5	5	5	5	5	5	50
VWC	5	5	5	5	5	5	5	5	5	5	50
RLH	5	5	5	5	5	5	5	5	5	5	50
TCA	4	4	4	4	4	4	4	4	4	4	40
WEA	12	12	12	12	12	12	12	12	12	12	120
										Total	450

Table 2-4

Number of Possible Predictors When Used in Cumulative Form

Elem.	Station ACY	NEL	PHL	SBY	MDT	DCA	PSB	BGM	RIC	IDL	Total
TOD	1	-	-	-	-	-	-	-	-	-	1
DOY	1	-	-	-	-	-	-	-	-	-	1
CIG	22	4	4	4	4	4	4	4	4	4	58
VIS	22	4	4	4	4	4	4	4	4	4	58
UWC	4	4	4	4	4	4	4	4	4	4	40
VWC	4	4	4	4	4	4	4	4	4	4	40
RLH	4	4	4	4	4	4	4	4	4	4	40
TCA	3	3	3	3	3	3	3	3	3	3	30
WEA	11	11	11	11	11	11	11	11	11	11	110
									Total		378

Table 2-5

Specifications and Limits of the Zero-One (Dummy) Variables Used in the Standard Dummy Form

Meteorological Element	Number of Dummies	Limits of Each Category
		Time of Day:
TOD	1	1. 0130 EST to 1330 EST
		Day of Year:
DOY	1	1. From 151 (31 May) to 300 (27 Oct)
		Ceiling Height, ft:
CIG*	5	1. $0 \leq CIG < 200$
		2. $200 \leq CIG < 500$
		3. $500 \leq CIG < 1000$
		4. $1000 \leq CIG < 3000$
		5. $3000 \leq CIG$
	19	(1) $CIG = 0$
		(2) $CIG = 100$
		(3) $CIG = 200$
		(4) $CIG = 300$
		(5) $CIG = 400$
		(6) $CIG = 500$
		(7) $CIG = 600$
		(8) $CIG = 700$
		(9) $CIG = 800$
		(10) $CIG = 900$
		(11) $1000 \leq CIG < 1500$
		(12) $1500 \leq CIG < 2000$
		(13) $2000 \leq CIG < 2500$
		(14) $2500 \leq CIG < 3000$
		(15) $3000 \leq CIG < 5000$
		(16) $5000 \leq CIG < 10,000$

*The predictand station (ACY) has ceiling and visibility dummied in two ways. Parentheses are used to identify the second set of dummies.

Table 2-5 - Continued

Meteorological Element	Number of Dummies	Limits of Each Category
		Ceiling Height, ft:
		(17) 10,000 ≤ CIG < 20,000
		(18) 20,000 ≤ CIG < Unlimited
		(19) CIG = Unlimited
		Visibility, mi:
VIS*	5	1. 0 ≤ VIS < 1/2
		2. 1/2 ≤ VIS < 1
		3. 1 ≤ VIS < 2
		4. 2 ≤ VIS < 3
		5. 3 ≤ VIS
	19	(1) 0 ≤ VIS < 1/16
		(2) 1/16 ≤ VIS < 1/8
		(3) 1/8 ≤ VIS < 1/4
		(4) 1/4 ≤ VIS < 1/2
		(5) 1/2 ≤ VIS < 3/4
		(6) 3/4 ≤ VIS < 1
		(7) 1 ≤ VIS < 1-1/2
		(8) 1-1/2 ≤ VIS < 2
		(9) 2 ≤ VIS < 3
		(10) 3 ≤ VIS < 4
		(11) 4 ≤ VIS < 5
		(12) 5 ≤ VIS < 6
		(13) 6 ≤ VIS < 7
		(14) 7 ≤ VIS < 8
		(15) 8 ≤ VIS < 9
		(16) 9 ≤ VIS < 10
		(17) 10 ≤ VIS < 11

*The predictand station (ACY) has ceiling and visibility dummies in two ways. Parentheses are used to identify the second set of dummies.

Table 2-5 - Continued

Meteorological Element	Number of Dummies		Limits of Each Category
			Visibility, mi:
		(18)	11 ≤ VIS < 15
		(19)	15 ≤ VIS
			"U" Wind Component (East-West Comp.):
UWC	5	1.	UWC ≤ -15
		2.	-15 < UWC ≤ - 5
		3.	- 5 < UWC ≤ 5
		4.	5 < UWC ≤ 15
		5.	15 < UWC
			"V" Wind Component (North-South Comp.):
VWC	5	1.	VWC ≤ -15
		2.	-15 < VWC ≤ - 5
		3.	- 5 < VWC ≤ 5
		4.	5 < VWC ≤ 15
		5.	15 < VWC
			Relative Humidity, per cent:
RLH	5	1.	RLH ≤ 50
		2.	50 < RLH ≤ 75
		3.	75 < RLH ≤ 90
		4.	90 < RLH ≤ 99
		5.	99 < RLH
			Total Cloud Amount:
TCA	4	1.	TCA < 1/10
		2.	1/10 ≤ TCA ≤ 5/10
		3.	5/10 < TCA ≤ 9/10
		4.	9/10 < TCA

Table 2-5 - Continued

Meteorological Element	Number of Dummies	Limits of Each Category	
		Weather Element:	Teletype Code:
WEA	12	1. No Weather	
		2. Rain	R-, R, R+
		3. Rain Showers	RW-, RW, RW+
		4. Drizzle or	L-, L, L+
		Freezing Drizzle	ZL-, ZL, ZL+
		5. Snow	S-, S, S+
		Snow Pellets	SP-, SP, SP+
		Snow Grains	SG-, SG, SG+
		Ice Crystals	IC-, IC, IC+
		6. Snow Showers	SW-, SW, SW+
		Snow Squalls	SQ-, SQ, SQ+
		7. Tornado	Tornado
		Thunderstorm	T, T+
		Squalls	Q-, Q, Q+
		Hail	A-, A, A+
		Hail Pellets	AP-, AP, AP+
		8. Freezing Rain	ZR-, ZR, ZR+
		Sleet	E-, E, E+
		Sleet Showers	EW-, EW, EW+
		9. Fog, Ice Fog	F, IF
		10. Ground Fog	GF
		11. Blowing Snow	BS
		Blowing Sand	BN
		Blowing Dust	BD
		12. Smoke, Haze, Dust	K, H, D

Table 2-6

Specifications and Limits of the Zero-One (Dummy) Variables
Used in the Cumulative Dummy Form

Meteorological Element	Number of Dummies	Limits of Each Category
		Time of Day
TOD	1	1. 0130 EST to 1330 EST
		Day of Year
DOY	1	1. From 151 (31 May) to 300 (27 Oct)
		Ceiling Height, ft:
CIG*	4	1. CIG < 200
		2. CIG < 500
		3. CIG < 1000
		4. CIG < 3000
	18	(1) CIG = 0
		(2) CIG ≤ 100
		(3) CIG ≤ 200
		(4) CIG ≤ 300
		(5) CIG ≤ 400
		(6) CIG ≤ 500
		(7) CIG ≤ 600
		(8) CIG ≤ 700
		(9) CIG ≤ 800
		(10) CIG ≤ 900
		(11) CIG < 1500
		(12) CIG < 2000
		(13) CIG < 2500
		(14) CIG < 3000
		(15) CIG < 5000

*The predictand station (ACY) has ceiling and visibility dummied in two ways. Parentheses are used to identify the second set of dummies.

Table 2-6 - Continued

Meteorological Element	Number of Dummies	Limits of Each Category
		Ceiling Height, ft:
		(16) CIG < 10,000
		(17) CIG < 20,000
		(18) CIG < Unlimited
		Visibility, mi:
VIS*	4	1. VIS < 1/2
		2. VIS < 1
		3. VIS < 2
		4. VIS < 3
	18	(1) VIS < 1/16
		(2) VIS < 1/8
		(3) VIS < 1/4
		(4) VIS < 1/2
		(5) VIS < 3/4
		(6) VIS < 1
		(7) VIS < 1-1/2
		(8) VIS < 2
		(9) VIS < 3
		(10) VIS < 4
		(11) VIS < 5
		(12) VIS < 6
		(13) VIS < 7
		(14) VIS < 8
		(15) VIS < 9
		(16) VIS < 10
		(17) VIS < 11
		(18) VIS < 15

*The predictand station (ACY) has ceiling and visibility dummied in two ways. Parentheses are used to identify the second set of dummies.

Table 2-6 - Continued

Meteorological Element	Number of Dummies	Limits of Each Category
		"U" Wind Component (East-West Comp.):
UWC	4	1. UWC $\leq$ -15
		2. UWC $\leq$ - 5
		3. UWC $\leq$ 5
		4. UWC $\leq$ 15
		"V" Wind Component (North-South Comp.):
VWC	4	1. VWC $\leq$ -15
		2. VWC $\leq$ - 5
		3. VWC $\leq$ 5
		4. VWC $\leq$ 15
		Relative Humidity, per cent:
RLH	4	1. RLH $\leq$ 50
		2. RLH $\leq$ 75
		3. RLH $\leq$ 90
		4. RLH $\leq$ 99
		Total Cloud Amount:
TCA	3	1. TCA $<$ 1/10
		2. TCA $\leq$ 5/10
		3. TCA $\leq$ 9/10

Table 2-6 - Continued

Meteorological Element	Number of Dummies	Limits of Each Category	
		Weather Element:	Teletype Code
WEA	11	1. No Weather	(See Table 2-5)
		2. Rain	
		3. Rain, Rain Showers	
		4. Rain, Rain Showers Freezing Rain, Sleet Sleet Showers	
		5. Drizzle Freezing Drizzle	
		6. Snow Snow Pellets Snow Grains Ice Crystals	
		7. Snow Snow Pellets Snow Grains Ice Crystals Snow Showers Snow Squalls	
		8. Tornado Thunderstorm Squalls Hail Hail Pellets	
		9. Fog, Ice Fog	
		10. Fog, Ice Fog Ground Fog	
		11. Smoke, Haze, Dust	

Table 2-7a

Selected Predictors and REEP Equations
ACY 2-hr Ceiling - Standard Dummies*

Selected Predictors				Coefficients of the REEP Equations				
Order	Station	Element	Dummy*	A_1	A_2	A_3	A_4	A_5
1	ACY	CIG	5	-.006	.013	-.016	-.349	.358
2	ACY	CIG	1	.294	.267	-.089	-.415	-.057
3	ACY	CIG	2	.027	.402	.001	-.395	-.035
4	PHL	CIG	5	.012	-.026	-.071	-.112	.198
5	ACY	CIG	3	.002	.081	.283	-.347	-.019
6	SBY	CIG	5	.000	-.069	-.053	.001	.121
7	NEL	TCA	4	.000	.001	.010	.044	-.055
8	ACY	CIG	(1)	.240	-.147	-.001	.008	-.100
9	NEL	CIG	3	-.064	-.050	.092	.044	-.022
10	ACY	VIS	5	-.036	-.011	-.023	.030	.041
11	DCA	CIG	2	.095	-.004	-.088	.029	-.032
12	NEL	CIG	4	-.058	-.048	-.024	.144	-.013
13	NEL	CIG	5	-.063	-.038	-.036	.063	.074
14	DCA	CIG	5	.006	-.015	-.024	-.039	.071
15	BGM	WEA	8	.146	-.104	-.056	-.024	.038
16	PHL	WEA	2	-.023	.063	.046	-.003	-.083
17	DCA	VIS	1	.148	-.072	-.051	-.022	-.003
18	SBY	CIG	4	-.007	-.068	-.058	.108	.025
19	PHL	UWC	2	-.000	.021	.016	.026	-.064
20	ACY	VIS	(5)	.103	-.035	-.011	-.055	-.002
		Additive Constant:		.089	.143	.221	.413	.134

*Refer to Table 2-5.

Table 2-7b

Selected Predictors and REEP Equations
ACY 2-hr Ceiling - Cumulative Dummies*

Order	Selected Predictors Station	Element	Dummy*	Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1	ACY	CIG	4	.009	-.012	-.005	.301	-.292
2	ACY	CIG	1	.232	-.109	-.072	-.036	-.015
3	ACY	CIG	2	.036	.313	-.272	-.051	-.026
4	PHL	CIG	4	-.011	.024	.073	.125	-.211
5	ACY	CIG	3	.010	.083	.269	-.401	.038
6	SBY	CIG	4	-.008	.001	-.006	.111	-.098
7	ACY	VIS	(5)	.098	-.104	-.003	.024	-.015
8	NEL	TCA	3	-.002	.001	-.008	-.053	.063
9	DCA	CIG	2	.103	-.033	-.074	.006	-.002
10	ACY	CIG	(1)	.208	-.129	.008	.009	-.097
11	NEL	CIG	3	-.000	-.007	.117	-.067	-.044
12	IDL	UWC	2	.009	.009	.028	.014	-.059
13	SBY	CIG	3	.000	.069	.075	-.115	-.029
14	ACY	CIG	(12)	-.002	-.006	.038	.127	-.157
15	ACY	RIH	4	.009	-.064	-.001	.060	-.003
16	DCA	CIG	4	-.007	.017	.023	.042	-.076
17	NEL	CIG	2	.057	.057	-.079	-.037	.002
18	SBY	VIS	2	.056	.017	-.075	.002	-.001
19	PHL	WEA	4	-.006	.032	.035	.017	-.078
		Additive Constant:		-.006	.061	.007	.001	.937

*Refer to Table 2-6.

Table 2-7c

Selected Predictors and REEP Equations
ACY 5-hr Ceiling - Standard Dummies*

Order	Selected Predictors Station	Element	Dummy*	Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1	ACY	CIG	5	-.003	-.002	-.006	-.154	.166
2	DCA	CIG	5	-.026	-.107	-.032	.018	.147
3	PHL	CIG	5	.001	-.024	-.045	-.070	.139
4	ACY	CIG	1	.183	.078	.092	-.229	-.124
5	PHL	UWC	2	-.000	.044	.043	.032	-.119
6	ACY	CIG	2	.045	.168	.048	-.266	.005
7	MDT	TCA	4	-.003	.013	.016	.052	-.078
8	SBY	CIG	5	-.004	.010	-.031	-.075	.100
9	DCA	CIG	4	-.028	-.102	.011	.105	.014
10	MDT	WEA	2	-.011	.042	.101	-.027	-.105
11	ACY	TOD	1	-.027	-.014	.002	.032	.007
12	ACY	CIG	3	.008	.055	.108	-.143	-.028
13	ACY	VIS	5	-.050	.017	-.005	-.018	.056
14	IDL	CIG	5	-.012	-.009	-.030	-.035	.086
15	SBY	WEA	9	.002	.068	.046	-.022	-.094
16	IDL	UWC	2	.005	.008	.040	.026	-.080
	Additive Constant:			.111	.117	.143	.330	.298

*Refer to Table 2-5.

Table 2-7d

Selected Predictors and REEP Equations
ACY 5-hr Ceiling - Cumulative Dummies*

Order	Selected Predictors Station	Element	Dummy*	Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1	ACY	CIG	(13)	.002	.004	.021	.156	-.182
2	DCA	CIG	4	-.002	.004	.040	.089	-.131
3	PHL	CIG	4	.000	.025	.044	.075	-.144
4	ACY	CIG	(3)	.122	-.110	.010	.003	-.025
5	PHL	UWC	2	-.000	.040	.051	.029	-.120
6	ACY	CIG	2	.024	.137	-.049	-.109	-.004
7	MDT	TCA	3	.004	-.014	-.015	-.051	.075
8	SBY	CIG	4	.005	-.012	.027	.079	-.100
9	IDL	RLH	2	-.012	-.004	-.020	-.034	.070
10	DCA	CIG	3	.026	.105	-.009	-.103	-.019
11	MDT	WEA	2	-.012	.038	.095	-.032	-.089
12	ACY	TOD	1	-.026	-.014	.003	.032	.005
13	ACY	CIG	3	.010	.053	.102	-.148	-.018
14	SBY	WEA	9	.001	.068	.044	-.026	-.087
15	IDL	UWC	2	.006	.013	.037	.025	-.081
16	ACY	VIS	4	.045	-.013	.002	.011	-.044
		Additive Constant:		.022	.018	.024	.074	.862

*Refer to Table 2-6.

Table 2-7e

Selected Predictors and REEP Equations
ACY 7-hr Ceiling - Standard Dummies*

Order	Selected Predictors Station	Element	Dummy*	Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1	PHL	CIG	5	.004	-.035	-.046	-.078	.155
2	DCA	CIG	5	-.010	-.012	-.040	-.061	.122
3	ACY	CIG	5	-.017	-.030	-.042	-.076	.165
4	PHL	UWC	2	.003	.025	.057	.042	-.132
5	MDT	TCA	4	-.001	.013	.020	.051	-.084
6	DCA	WEA	2	-.003	.056	.125	-.038	-.139
7	ACY	VIS	5	-.044	-.001	.007	.020	.019
8	PSB	WEA	8	-.020	.248	-.095	-.047	-.086
9	ACY	TOD	1	-.028	-.019	-.009	.024	.031
10	RIC	RIH	4	-.005	.014	.032	.040	-.081
11	NEL	WEA	9	.018	.019	.062	.013	-.112
12	IDL	UWC	2	.011	.017	.035	.020	-.082
13	ACY	CIG	2	.025	.101	.003	-.132	.003
14	ACY	CIG	(1)	.155	-.010	-.015	.006	-.135
15	DCA	CIG	3	-.004	.093	.007	-.051	-.045
16	NEL	UWC	4	-.005	-.010	-.011	-.028	.054
	Additive Constant:			.087	.087	.124	.212	.491

*Refer to Table 2-5.

Table 2-7f

Selected Predictors and REEP Equations
ACY 7-hr Ceiling - Cumulative Dummies*

Selected Predictors				Coefficients of the REEP Equations				
Order	Station	Element	Dummy*	A_1	A_2	A_3	A_4	A_5
1	PHL	CIG	4	-.004	.025	.047	.082	-.151
2	DCA	CIG	4	.001	-.002	.039	.084	-.123
3	ACY	CIG	(13)	.008	.018	.014	.097	-.137
4	PHL	UWC	2	.007	.033	.063	.032	-.135
5	MDT	TCA	3	-.000	-.010	-.014	-.049	.073
6	ACY	VIS	2	.091	-.040	-.052	.032	-.030
7	DCA	WEA	4	-.017	.022	.094	-.020	-.079
8	IDL	RLH	2	-.008	.000	-.017	-.040	.066
9	DCA	CIG	3	.016	.098	.004	-.093	-.025
10	ACY	TCD	1	-.027	-.018	-.004	.027	.023
11	NEL	UWC	3	.006	.008	.012	.030	-.057
12	SBY	WEA	9	.010	.044	.019	.039	-.113
13	ACY	CIG	(9)	.037	.064	.072	-.134	-.038
14	IDL	UWC	2	.011	.017	.030	.020	-.078
15	PHL	WEA	4	-.000	.056	.029	-.027	-.057
		Additive Constant:		.021	.011	.021	.069	.878

*Refer to Table 2-6.

Table 2-8a

Contingency Tables Showing Predicted-Versus-Actual Groups for Persistence, REEP with Standard Dummies, and REEP with Cumulative Dummies

2-hr Ceiling (Independent Sample)

		Actual					
	Group	1	2	3	4	5	Total
PERSISTENCE	1	26	4	3	0	5	38
	2	8	48	22	6	6	90
	3	4	10	57	24	13	108
	4	2	4	24	72	66	168
	5	4	9	16	55	1556	1640
	Total	44	75	122	157	1646	2044
		Hits = 1759			$\bar{P}$ = .2063		
	$\bar{P}_g$	.0128	.0241	.0407	.0590	.0696	

		Actual					
	Group	1	2	3	4	5	Total
REEP with Standard Dummies	1	19	2	2	0	4	27
	2	11	46	21	5	2	85
	3	3	8	52	21	10	94
	4	2	4	21	58	31	116
	5	9	15	26	73	1599	1722
	Total	44	75	122	157	1646	2044
		Hits = 1774			$\bar{P}$ = .1878		
	$\bar{P}_g$	.0138	.0237	.0369	.0558	.0575	

		Actual					
	Group	1	2	3	4	5	Total
REEP with Cumulative Dummies	1	25	3	3	0	4	35
	2	6	46	21	5	2	80
	3	3	8	52	19	10	92
	4	1	4	24	56	26	111
	5	9	14	22	77	1604	1726
	Total	44	75	122	157	1646	2044
		Hits = 1783			$\bar{P}$ = .1884		
	$\bar{P}_g$	.0134	.0236	.0367	.0571	.0576	

Table 2-8b

Contingency Tables Showing Predicted-Versus-Actual Groups for Persistence, REEP with Standard Dummies, and REEP with Cumulative Dummies

5-hr Ceiling (Independent Sample)

		Actual					
	Group	1	2	3	4	5	Total
PERSISTENCE	1	11	11	6	2	8	38
	2	7	34	18	15	16	90
	3	3	13	38	28	26	108
	4	2	11	29	47	79	168
	5	9	15	31	82	1503	1640
	Total	32	84	122	174	1632	2044
		Hits = 1633			$\bar{P}$ = .2689		
	$\bar{P}_g$	.0138	.0326	.0489	.0723	.1013	

		Actual					
	Group	1	2	3	4	5	Total
REEP with Standard Dummies	1	1	1	1	0	0	3
	2	3	24	7	6	2	42
	3	7	20	32	17	8	84
	4	0	10	24	23	11	68
	5	21	29	58	128	1611	1847
	Total	32	84	122	174	1632	2044
		Hits = 1691			$\bar{P}$ = .2375		
	$\bar{P}_g$	.0136	.0310	.0460	.0687	.0782	

		Actual					
	Group	1	2	3	4	5	Total
REEP with Cumulative Dummies	1	0	0	0	1	0	1
	2	1	19	8	4	3	35
	3	9	23	31	17	5	85
	4	0	9	21	18	9	57
	5	22	33	62	134	1615	1866
	Total	32	84	122	174	1632	2044
		Hits = 1683			$\bar{P}$ = .2367		
	$\bar{P}_g$	.0137	.0310	.0459	.0685	.0776	

Table 2-8c

Contingency Tables Showing Predicted-Versus-Actual Groups for Persistence, REEP with Standard Dummies, and REEP with Cumulative Dummies

7-hr Ceiling (Independent Sample)

		Actual					
	Group	1	2	3	4	5	Total
PERSISTENCE	1	8	12	8	4	6	38
	2	12	22	19	19	18	90
	3	7	9	38	26	28	108
	4	2	12	33	37	84	168
	5	12	24	49	72	1483	1640
	Total	41	79	147	158	1619	2044
		Hits = 1588			$\bar{P}$ = .2899		
	$\bar{P}_g$	.0183	.0342	.0598	.0671	.1105	

		Actual					
	Group	1	2	3	4	5	Total
REEP with Standard Dummies	1	0	1	0	0	0	1
	2	1	8	5	0	3	17
	3	9	22	32	9	7	79
	4	7	6	18	12	5	48
	5	24	42	92	137	1604	1899
	Total	41	79	147	158	1619	2044
		Hits = 1656			$\bar{P}$ = .2617		
	$\bar{P}_g$	.0184	.0322	.0547	.0668	.0896	

		Actual					
	Group	1	2	3	4	5	Total
REEP with Cumulative Dummies	1	0	0	0	0	0	0
	2	5	9	6	1	0	21
	3	7	19	28	10	4	68
	4	4	7	21	10	9	51
	5	25	44	92	137	1606	1904
	Total	41	79	147	158	1619	2044
		Hits = 1653			$\bar{P}$ = .2621		
	$\bar{P}_g$	.0182	.0323	.0556	.0674	.0885	

Table 2-9

Comparison Between Persistence, REEP with Standard Dummies, and REEP with Cumulative Dummies for the Independent Sample (2044 Observations)

2-hr Ceiling		
Hits	Persistence	1759
	Standard	1774
	Cumulative	1783*
$\bar{P}$ Score	Persistence	.2063
	Standard	.1878*
	Cumulative	.1884
5-hr Ceiling		
Hits	Persistence	1633
	Standard	1691*
	Cumulative	1683
$\bar{P}$ Score	Persistence	.2689
	Standard	.2375
	Cumulative	.2367*
7-hr Ceiling		
Hits	Persistence	1588
	Standard	1656*
	Cumulative	1653
$\bar{P}$ Score	Persistence	.2899
	Standard	.2617*
	Cumulative	.2621

Asterisks are used to denote superiority.

EXPERIMENT 3. Nullifying the Effect of High Ceilings in Predictor Selection.

Purpose

To determine the effect on the forecast of assuming a zero degree of freedom for category 5 of the ceiling predictand.

Procedure

Forecasts were made for ceiling at Atlantic City 3 hours and 7 hours in advance. The degrees of freedom specified for categories 1 through 4 of the ceiling predictand was specified as 2652, the number normally used for the Atlantic City developmental sample size of 7956, or one-third of the total developmental sample. The degrees of freedom for category 5 of the ceiling predictand was specified as zero. This in effect eliminated the possibility of having category 5 affect the selection of any predictor.

Data Sample

The ten years of data from 1 January 1949 to 31 December 1958, consisting of 10,000 observations selected at random was used in this experiment. The developmental sample consisted of 7956 observations from the first eight years and the verification or dependent sample of 2044 was from the last two years of the data sample.

Predictands

The ceiling at time intervals of 3 and 7 hours in advance were the predictands for this experiment. The operationally significant categories for which the forecasts were made are shown in Table 3-1. The specifications are shown in Table 3-2.

Predictors

The set of 423 dummy predictors used in Experiment No. 1 for the 13-station Atlantic City network was used for this experiment (see Figure 1-1 and Tables 1-3 and 1-4).

Results

The predictors selected by REEP for each of the predictands are given in Tables 3-3a and 3-3b, as well as the coefficients in the REEP regression equation for each group of the predictand. Contingency tables showing the forecast-versus-observed frequencies for REEP with zero degree of freedom in category 5 and for persistence are shown in Tables 3-4a and 3-4b. The number of correct forecasts (hits), the Brier and Allen $\bar{P}$ scores, and the individual group contributions to the overall $\bar{P}$ score are shown.

Table 3-5 summarizes the hits and $\bar{P}$ scores and compares them with the results using the regular degree of freedom for all categories of the predictand as determined by Experiment No. 1.

Conclusions

The results in this experiment are the same or slightly worse than those using the regular degrees of freedom, both with regard to $\bar{P}$ score and hits as well as with regard to better predicting the low categories. It is therefore concluded that modifying the degrees of freedom within the higher categories does not appreciably change the results from those obtained using the regular degrees of freedom. The lack of success in this effort to force the selection of predictors which contribute to improving the forecasts of low ceiling conditions suggests that the regular selection method is performing this function satisfactorily.

Table 3-1

Definition of Predictand Groups for Atlantic City Municipal Airport

Group Number	Ceiling Intervals
1	0 ft $\leq$ Ceiling $<$ 200 ft
2	200 ft $\leq$ Ceiling $<$ 500 ft
3	500 ft $\leq$ Ceiling $<$ 1000 ft
4	1000 ft $\leq$ Ceiling $<$ 3000 ft
5	3000 ft $\leq$ Ceiling

Table 3-2

Specification of Predictands
Atlantic City Municipal Airport

	Predictand 1	Predictand 2
Predictand variable	CIG	CIG
Forecast interval--hours (H)	3	7
Number of Groups (G)	5	5
Observations in Group 1 (n_1)	100	135
Observations in Group 2 (n_2)	281	272
Observations in Group 3 (n_3)	385	420
Observations in Group 4 (n_4)	660	619
Observations in Group 5 (n_5)	6530	6510
Total dependent sample size (N)	7956	7956
Number of independent observations in the dependent sample (n)		
Group 1	2652	2652
Group 2	2652	2652
Group 3	2652	2652
Group 4	2652	2652
Group 5	0	0
Total independent sample size (M)	2044	2044
Number of available predictors (P)	423	423

Table 3-3a

Selected Predictors and REEP Equations
Atlantic City 3-hr Ceiling Forecast

(Zero Degrees Freedom, Category 5)

Predictors Selected				Coefficients of the REEP Equations				
Order	Station	Element	Dummy	A_1	A_2	A_3	A_4	A_5
1	ACY	CIG	2	.057	.329	.025	-.075	-.336
2	ACY	CIG	1	.396	.202	.001	-.146	-.453
3	ACY	CIG	4	.008	.016	.027	.199	-.250
4	ACY	CIG	3	.007	.082	.214	-.013	-.291
5	PHL	CIG	5	.008	-.036	-.082	-.143	.253
6	DCA	RLH	3	.012	.079	.017	-.033	-.075
7	SBY	CIG	4	-.009	-.079	-.033	.089	.032
8	SBY	WEA	2	-.009	.084	-.002	.053	-.124
9	NEL	CIG	3	-.008	-.010	.098	.008	-.088
10	ACY	VIS	5	-.047	.013	-.006	-.001	.041
11	ACY	CIG	(1)	-.176	.064	-.069	.049	.132
12	NEL	WEA	2	-.025	.021	.087	-.046	-.036
13	ACY	TCA	2	.004	-.008	.016	.054	-.066
14	SBY	CIG	5	-.004	-.071	-.052	-.023	.150
15	SBY	VIS	2	.085	.007	-.037	-.031	-.024
16	ACY	CIG	(6)	-.009	-.012	-.041	.164	-.102
17	NEL	CIG	4	-.003	-.017	.012	.082	-.074
		Additive Constant:		.045	.094	.141	.180	.539

Table 3-3b

Selected Predictors and REEP Equations
Atlantic City 7-hr Ceiling Forecast

(Zero Degrees Freedom, Category 5)

Predictors Selected				Coefficients of the REEP Equations				
Order	Station	Element	Dummy	A_1	A_2	A_3	A_4	A_5
1	PHL	CIG	5	-.009	-.048	-.060	-.079	.195
2	DCA	WEA	2	-.003	.055	.126	-.041	-.136
3	ACY	SCL	1	.112	-.025	-.016	.006	-.077
4	ACY	CIG	2	.017	.103	.019	-.124	-.016
5	SBY	CIG	5	-.010	-.000	-.036	-.077	.123
6	DCA	CIG	3	.005	.102	.055	.025	-.187
7	PHL	TCA	2	-.003	.013	.023	.055	-.087
8	PSB	WEA	8	-.024	.243	-.065	-.030	-.105
9	ACY	TOD	1	-.026	-.019	-.009	.024	.030
10	NEL	WND	2	-.005	.038	.064	.025	-.122
11	NEL	WEA	9	.028	.021	.071	.015	-.135
12	DCA	CIG	4	.000	.012	.047	.085	-.145
13	PSB	CIG	1	.014	.098	-.017	-.023	-.072
14	BGM	RLH	3	-.004	-.028	.008	.065	-.041
15	PHL	WND	3	.029	.022	.037	.042	-.130
16	IDL	CIG	3	.010	.072	-.004	-.015	-.063
		Additive Constant:		.037	.056	.097	.157	.654

Table 3-4a

Contingency Tables Showing Predictand-Versus-Actual Groups for Persistence and REEP Forecast With Zero Degree of Freedom in Category 5

3-hr Ceiling (Verification Sample)

		Actual					
	Group	1	2	3	4	5	Total
Persistence	1	21	8	4	0	5	38
	2	13	33	25	10	9	90
	3	2	14	50	21	21	108
	4	1	8	21	68	70	168
	5	7	10	27	66	1530	1640
	Total	44	73	127	165	1635	2044
		Hits = 1702			$\bar{P}$ = .2385		
	P_g	.0156	.0284	.0462	.0636	.0847	

		Actual					
	Group	1	2	3	4	5	Total
REEP With Zero Degree of Freedom In Category 5	1	7	4	1	0	1	13
	2	18	34	24	8	7	91
	3	0	10	40	15	7	72
	4	1	7	19	42	18	87
	5	18	18	43	100	1602	1781
	Total	44	73	127	165	1635	2044
		Hits = 1725			$\bar{P}$ = .2131		
	P_g	.0163	.0275	.0429	.0591	.0673	

Table 3-4b

Contingency Tables Showing Predictand-Versus-Actual Groups for Persistence and REEP Forecast with Zero Degree of Freedom in Category 5

7-hr Ceiling (Verification Sample)

		Actual					
	Group	1	2	3	4	5	Total
Persistence	1	8	12	8	4	6	38
	2	12	22	19	19	18	90
	3	7	9	38	26	28	108
	4	2	12	33	37	84	168
	5	12	24	49	72	1483	1640
	Total	41	79	147	158	1619	2044
		Hits = 1588			$\bar{P}$ = .2899		
	P_g	.0183	.0342	.0598	.0671	.1105	

		Actual					
	Group	1	2	3	4	5	Total
REEP With Zero Degree of Freedom in Category 5	1	0	0	0	0	0	0
	2	3	11	12	3	4	33
	3	5	18	28	10	6	67
	4	5	5	21	13	8	52
	5	28	45	86	132	1601	1892
	Total	41	79	147	158	1619	2044
		Hits = 1653			$\bar{P}$ = .2638		
	P_g	.0181	.0325	.0546	.0667	.0920	

Table 3-5

Comparison Between Persistence, REEP with Zero Degrees of Freedom in Category 5, REEP with the Regular Degrees of Freedom in All Categories (See Experiment No. 1)

Atlantic City - Independent Sample

Ceiling: 3-hr Forecast

Hits	Persistence	1702
	REEP - Zero degrees freedom in Category 5	1725
	REEP - Regular degrees freedom in all categories	1728*
$\bar{P}$ Score	Persistence	.2385
	REEP - Zero degrees freedom in Category 5	.2131*
	REEP - Regular degrees freedom in all categories	.2135

Ceiling: 7-hr Forecast

Hits	Persistence	1588
	REEP - Zero degrees freedom in Category 5	1653
	REEP - Regular degrees freedom in all categories	1665*
$\bar{P}$ Score	Persistence	.2899
	REEP - Zero degrees freedom in Category 5	.2638
	REEP - Regular degrees freedom in all categories	.2600*

*Asterisks are used to denote superiority.

EXPERIMENT 4. Ceiling Prediction Using Persistence with One and Two Hour Lagged Persistence.

Purpose

A test on the predictability of persistence when time variations in persistence are made available.

Predictands

Predictands chosen for this experiment were ceilings at 1, 2, and 3 hours in advance. The operationally significant categories for which the forecasts were made are shown in Table 4-1. The specifications for the predictands are shown in Table 4-2.

Predictors

All the possible combinations of the five categories of ceiling for Atlantic City at the three time periods: 1) time of observation, 2) time of observation minus one hour, and 3) time of observation minus two hours, were used in combinations of logical "ANDS" and logical "ORS" to form 400 predictors. These 400 predictors plus the 15 ceiling dummies for the three different time periods gave a total of 435 predictors, made up only of ceiling dummies for only one station--Atlantic City. The intervals for the five categories of ceiling used as predictors are the same as those for the predictands shown in Table 4-1.

Data Sample

Standard hourly airways surface observations covering the ten year period, 1 January 1949 to 31 December 1958, furnished the data for this study. The developmental sample of 7956 observations is a random selection from the first eight years and the verification sample of 2044 observations is a random selection from the last two years. Note: This sample is not the same as in previous experiments because lagged predictors were involved.

Results

The predictors selected by REEP for each of the three predictands are given in Tables 4-3a to 4-3c. Also shown are the REEP regression equation coefficients for each group of the predictand.

Contingency tables showing the forecast-versus-observed frequencies for each of the three forecasts, and comparing the results using REEP to the results using persistence (without lags) is shown in Tables 4-4a to 4-4c. The number of correct forecasts (hits) and the Brier and Allen $\bar{P}$ scores as well as the individual group contributions to the overall $\bar{P}$ score are shown below each contingency table. The number of hits and the overall $\bar{P}$ scores are summarized in Table 4-5.

Conclusions

The 415 possible predictors exhaust the ways in which persistence with lags could improve the forecast. Contrary to common belief the results of this experiment show there is little improvement over persistence when lagged persistence is made available for selection.

Table 4-1

Definition of Predictand Groups for Atlantic City Municipal Airport

Group Number	Ceiling Intervals
1	0 ft ≤ Ceiling < 200 ft
2	200 ft ≤ Ceiling < 500 ft
3	500 ft ≤ Ceiling < 1000 ft
4	1000 ft ≤ Ceiling < 3000 ft
5	3000 ft ≤ Ceiling

Table 4-2

Specifications of Predictands

Atlantic City Municipal Airport

	Predictand 1	2	3
Predictand variable	CIG	CIG	CIG
Forecast interval--hours (H)	1	2	3
Number of Groups (G)	5	5	5
Observations in Group 1 (n_1)	125	128	133
Observations in Group 2 (n_2)	273	259	272
Observations in Group 3 (n_3)	403	426	418
Observations in Group 4 (n_4)	655	631	620
Observations in Group 5 (n_5)	6500	6512	6513
Total dependent sample size (N)	7956	7956	7956
Number of independent observations in the dependent sample (n)	2652	2652	2652
Total independent sample size (M)	2044	2044	2044
Number of available predictors (P)	415	415	415

Table 4-3a

Selected Predictors and REEP Equations
Atlantic City 1-hr Ceiling Forecast (Experiment 4)

Order	Predictor Selected (see note) Sta.	Elem.	Predictor Ceiling Category at time: t_0		t_{-1}		t_{-2}	Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1	ACY	CIG	5					.005	-.011	-.084	-.544	.634
2	ACY	CIG	1					.767	.120	-.073	-.630	-.184
3	ACY	CIG	2					.124	.526	.049	-.588	-.110
4	ACY	CIG	3					.009	.115	.460	-.478	-.106
5	ACY	CIG				*	5	-.005	-.008	-.013	-.094	.120
6	ACY	CIG	2	*	2	*	2	-.235	.352	-.154	.010	.027
7	ACY	CIG			3	*	2	-.125	.156	.025	-.047	-.009
8	ACY	CIG	2			*	2	.186	-.165	.065	-.031	-.054
9	ACY	CIG	2			+	3	-.008	-.012	.086	-.017	-.049
			Additive Constant:					.003	.021	.102	.657	.217

Note: t_0 means at time of observation.

t_{-1} means one hour before time of observation.

t_{-2} means two hours before time of observation.

* means logical "AND"

\+ means logical "OR"

See Table 4-1 for limits of each category (group).

Table 4-3b

Selected Predictors and REEP Equations
Atlantic City 2-hr Ceiling Forecast (Experiment 4)

Order	Predictor Selected (see note) Sta.	Elem.	Predictor Ceiling Category at time: t_0	t_{-1}	t_{-2}	Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1	ACY	CIG	5			.001	-.024	-.071	-.380	.475
2	ACY	CIG	1			.545	.189	-.110	-.471	-.153
3	ACY	CIG	2			.117	.349	.168	-.411	-.223
4	ACY	CIG	3			.045	.013	.321	-.238	-.141
5	ACY	CIG		5	* 5	-.006	-.016	-.032	-.138	.192
6	ACY	CIG	1 *	1	* 2	.334	-.058	-.099	-.051	-.127
7	ACY	CIG	5 *	2		.339	-.134	-.037	-.074	-.094
8	ACY	CIG	3 +	2		-.013	.118	-.011	-.036	-.008
9	ACY	CIG	4 *	5	* 2	-.008	-.044	.185	-.356	.223
10	ACY	CIG	1 +	3	+ 1	-.021	- 042	.097	-.019	-.015
			Additive Constant:			.008	.044	.111	.542	.295

Note: t_0 means at time of observation

t_{-1} means one hour before time of observation

t_{-2} means two hours before time of observation

* means logical "AND"

\+ means logical "OR"

See Table 4-1 for limits of each category (group).

Table 4-3c

Selected Predictors and REEP Equations
Atlantic City 3-hr Ceiling Forecast (Experiment 4)

Order	Predictor Selected (see note) Sta.	Elem.	Predictor Ceiling Category at time: t_0	t_{-1}	t_{-2}	Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1	ACY	CIG	5			.004	-.018	-.074	-.285	.374
2	ACY	CIG	1			.389	.173	-.006	-.338	-.217
3	ACY	CIG	2			.111	.346	.167	-.320	-.304
4	ACY	CIG	3			.008	.105	.308	-.203	-.219
5	ACY	CIG		5 *	5	-.016	-.032	-.049	-.099	.196
6	ACY	CIG	2	*	1	-.033	-.246	.045	.186	.048
7	ACY	CIG		1 *	2	.175	.026	-.082	-.066	-.054
			Additive Constant:			.016	.056	.132	.415	.381

Note: t_0 means at time of observation

t_{-1} means one hour before time of observation

t_{-2} means two hours before time of observation

* means logical "AND"

\+ means logical "OR"

See Table 4-1 for limits of each category (group).

Table 4-4a

Contingency Tables Showing Predictand-Versus-Actual Groups for Persistence and REEP Forecast Using Persistence with One and Two-Hour Lagged Persistence of Ceiling

Atlantic City 1-hr Ceiling (Verification Sample)

		Actual					
	Group	1	2	3	4	5	Total
Persistence	1	32	7	2	0	3	44
	2	2	51	14	2	4	73
	3	1	15	80	22	9	127
	4	0	2	15	113	35	165
	5	3	5	8	33	1586	1635
	Total	38	80	119	170	1637	2044

Hits = 1862 $\quad \bar{P} = .1469$

P_g	.0073	.0204	.0316	.0433	.0442

		Actual					
	Group	1	2	3	4	5	Total
REEP With Lagged Persistence	1	32	7	2	0	3	44
	2	2	51	14	2	4	73
	3	1	15	80	22	9	127
	4	0	2	15	113	35	165
	5	3	5	8	33	1586	1635
	Total	38	80	119	170	1637	2044

Hits = 1862 $\quad \bar{P} = .1458$

P_g	.0075	.0207	.0318	.0427	.0432

Table 4-4b

Contingency Tables Showing Predictand-Versus-Actual Groups for Persistence and REEP Forecast Using Persistence With One and Two-Hour Lagged Persistence of Ceiling

Atlantic City 2-hr Ceiling (Verification Sample)

		Actual					
	Group	1	2	3	4	5	Total
Persistence	1	24	12	4	1	3	44
	2	2	41	18	4	8	73
	3	2	24	57	33	11	127
	4	1	0	21	89	54	165
	5	3	7	22	47	1556	1635
	Total	32	84	122	174	1632	2044
		Hits = 1767			$\bar{P}$ = .2016		
	P_g	.0095	.0267	.0436	.0571	.0647	

		Actual					
	Group	1	2	3	4	5	Total
REEP With Lagged Persistence	1	24	12	4	1	3	44
	2	2	32	13	3	5	55
	3	2	33	62	34	14	145
	4	1	0	14	70	41	126
	5	3	7	29	66	1569	1674
	Total	32	84	122	174	1632	2044
		Hits = 1757			$\bar{P}$ = .2009		
	P_g	.0102	.0267	.0429	.0572	.0639	

Table 4-4c

Contingency Tables Showing Predictand-Versus-Actual Groups for Persistence and REEP Forecast Using Persistence With One and Two-Hour Lagged Persistence of Ceiling

Atlantic City 3-hr Ceiling (Verification Sample)

		Actual					
	Group	1	2	3	4	5	Total
Persistence	1	20	13	5	1	5	44
	2	6	31	20	5	11	73
	3	4	26	53	32	12	127
	4	1	2	28	64	70	165
	5	5	11	32	61	1526	1635
	Total	36	83	138	163	1624	2044

Hits = 1694 $\quad \bar{P} = .2392$

P_g	.0130	.0301	.0514	.0626	.0822

		Actual					
	Group	1	2	3	4	5	Total
REEP With Lagged Persistence	1	20	13	5	1	5	44
	2	5	26	16	5	11	63
	3	5	31	57	32	12	137
	4	1	2	22	53	56	134
	5	5	11	38	72	1540	1666
	Total	36	83	138	163	1624	2044

Hits = 1696 $\quad \bar{P} = .2375$

P_g	.0132	.0306	.0512	.0623	.0802

Table 4-5

Comparison Between Persistence and REEP with Lagged Persistence of Ceiling

Ceiling: 1-hr Forecast		
Hits	Persistence	1862
	REEP with Lagged Persistence of Ceiling	1862
$\bar{P}$ Score	Persistence	.1469
	REEP with Lagged Persistence of Ceiling	.1458*
Ceiling: 2-hr Forecast		
Hits	Persistence	1767*
	REEP with Lagged Persistence of Ceiling	1757
$\bar{P}$ Score	Persistence	.2016
	REEP with Lagged Persistence of Ceiling	.2009*
Ceiling: 3-hr Forecast		
Hits	Persistence	1694
	REEP with Lagged Persistence of Ceiling	1696*
$\bar{P}$ Score	Persistence	.2392
	REEP with Lagged Persistence of Ceiling	.2375*

Asterisks are used to denote superiority.

EXPERIMENT 5. Multiple Predictands Processed Simultaneously.

Purpose

To investigate the selection of predictors for more than one predictand at a time in an attempt to reduce the number of analysis runs ordinarily required.

Procedure

Four REEP selection runs were performed using the following pairs of predictands as the predictand:

1) Ceiling plus 3 hours and visibility plus 3 hours
2) Ceiling plus 3 hours and ceiling plus 7 hours
3) Visibility plus 3 hours and visibility plus 7 hours
4) Ceiling plus 7 hours and visibility plus 7 hours

Predictors consisted of dummy (zero-one) variables of seven meteorological elements from a thirteen-station network (see Figure 1-1). The predictors selected from those runs were compared with those selected when the predictands were run singly. This comparison is shown in Table 5-1. The predictors selected when the predictands were run singly are indicated by an x under the appropriate ceiling or visibility predictand.

Predictands

The five operationally significant categories of ceiling and visibility for the predictands are shown in Table 5-2. The specifications of the predictands are listed in Table 5-3.

Predictors

The 423 predictors used in this experiment are the same as those used in Experiment No. 1 (see Tables 1-3 and 1-4).

Data Sample

Standard hourly airways surface observations covering the ten year period, 1 January 1949 to 31 December 1958, constituted the data used in this study. The developmental sample of 7956 observations is a random selection from the first eight years and the verification sample of 2044 observations is a random selection from the last two years.

Results

Contingency tables (Tables 5-4a and 5-4b) were formed to show the forecast-versus-observed frequencies for the 3-hr predictions of ceiling and visibility using the predictors selected by REEP when the predictand was a single predictand and also when the predictand was a pair of predictands. The number of correct forecasts (hits) and the Brier and Allen $\bar{P}$ scores as well as the individual group contributions to the overall $\bar{P}$ score are given below each contingency table.

Conclusions

There is sufficient redundancy in the predictands such that selection of predictors may be made on more than one predictand at a time. This applies even though they differ with regard to meteorological element or with regard to time leads. These conclusions were drawn from the fact that the predictors selected, when the predictands were processed singly, are all essentially among those selected when they are considered jointly. The one direct comparison for which results on independent data were obtained (Ceiling 3 hrs and Visibility 3 hrs) confirms this surmise.

Table 5-1

Predictors Selected When the Predictand is a Set

(Columns at left of each list show those predictors also selected when selection was run singly for the indicated predictand.)

CIG-3, VIS-3			CIG-3, CIG-7			VIS-3, VIS-7			CIG-7, VIS-7		
CIG 3	VIS 3	Predictors	CIG 3	VIS 7	Predictors	VIS 3	VIS 3	Predictors	CIG 7	VIS 7	Predictors
X		ACY CIG-5	X	X	ACY CIG-5	X		ACY VIS-5	X	X	PHL CIG-5
X	X	ACY VIS-5	X		ACY CIG-1	X		ACY VIS-1	X		DCA CIG-5
X		PHL CIG-5	X	X	PHL CIG-5	X		ACY RLH-3	X		ACY CIG-5
X		ACY CIG-1	X	X	ACY CIG-2	X	X	ACY VIS-2	X	X	PHL WND-3
X		ACY CIG-2	X	X	DCA CIG-5		X	PHL CIG-5		X	DCA VIS-5
X		DCA CIG-5	X		ACY CIG-3		X	ACY TOD(02-13)	X	X	ACY TOD(02-13)
X		ACY CIG-3		X	PHL WND-3		X	DCA VIS-5	X		IDL WND-2
	X	ACY VIS-1	X	X	IDL WND-2			SBY RLH-3			ACY WEA-1
		SBY RLH-3	X		NEL TCA-2		X	PHL WEA-1	X	X	DCA WEA-2
	X	ACY VIS-2	X		SBY CIG-5	X		PHL CIG-1	X		ACY SCL-1
X		NEL TCA-2		X	DCA WEA-2	X		NEL WEA-9			MDT TCA-2
X		SBY CIG-5	X		ACY VIS-5			NEL WND-3	X		PSB WEA-8
	X	ACY WEA-1		X	PSB WEA-8		X	DCA WEA-2	X		ACY CIG-2
X		DCA RLH-3	X		SBY CIG-4	X		SBY VIS-5	X		NEL WEA-9
X		NEL CIG-3		X	ACY TOD(02-13)	X		MDT VIS-2	X		SBY WEA-9
	X	NEL WEA-9	X		ACY CIG(1)		X	IDL VIS-3		X	IDL VIS-3
X		SBY WEA-2	X		NEL CIG-3	X		ORF RLH-3			PHL WND-2
	X	PHL CIG-1		X	PSB RLH-3			NEL WEA-12			NEL WEA-12
X		SWF WND-3	X		SBY VIS-2						
X		ACY CIG(1)	X		DCA RLH-3						
X		IDL WND-2	X		NEL WEA-2						
X		ACY CIG(6)	X		ACY CIG(6)						
X		SBY CIG-4									
	X	ORF RLH-3									
X	X	NEL WEA-2									
	X	MDT VIS-2									

Table 5-2

Definition of Predictand Groups for Atlantic City Municipal Airport

Group Number	Ceiling Intervals
1	0 ft ≤ Ceiling < 200 ft
2	200 ft ≤ Ceiling < 500 ft
3	500 ft ≤ Ceiling < 1000 ft
4	1000 ft ≤ Ceiling < 3000 ft
5	3000 ft ≤ Ceiling

Group Number	Visibility Intervals
1	0 mi < Visibility < 1/2 mi
2	1/2 mi < Visibility < 1 mi
3	1 mi < Visibility < 2 mi
4	2 mi < Visibility < 3 mi
5	3 mi < Visibility

Table 5-3

Specification of Predictands
Atlantic City Municipal Airport

	Predictand 1	Predictand 2
Predictand variable	CIG	VIS
Forecast interval--hours (H)	3	3
Number of Groups (G)	5	5
Observations in Group 1 (n_1)	100	153
Observations in Group 2 (n_2)	281	125
Observations in Group 3 (n_3)	385	200
Observations in Group 4 (n_4)	660	286
Observations in Group 5 (n_5)	6530	7192
Total dependent sample size (N)	7956	7956
Number of independent observations in the dependent sample (n)	2652	2652
Total independent sample size (M)	2044	2044
Number of available predictors (P)	423	423

Table 5-4a

Contingency Tables Showing Predicted-Versus-Actual Groups Using Predictors Selected when the Predictand was a Single Predictand and Using Predictors Selected when the Predictand was a Pair of Predictands

Atlantic City 3-hr Ceiling (Independent Sample)

		Actual					
	Group	1	2	3	4	5	Total
Using Predictors Selected When Predictand was Single	1	6	4	1	0	1	12
	2	19	34	25	8	4	90
	3	0	11	42	14	5	72
	4	1	6	19	41	20	87
	5	18	18	40	102	1605	1783
	Total	44	73	127	165	1635	2044

Hits = 1728 $\quad \bar{P} = .2135$

P_g	.0163	.0277	.0423	.0605	.0666

		Actual					
	Group	1	2	3	4	5	Total
Using Predictors Selected When Predictand was a Pair	1	6	3	1	0	2	12
	2	19	35	25	8	4	91
	3	0	13	40	14	5	72
	4	1	7	19	41	22	90
	5	18	15	42	102	1602	1779
	Total	44	73	127	165	1635	2044

Hits = 1724 $\quad \bar{P} = .2133$

P_g	.0165	.0279	.0421	.0605	.0663

Table 5-4b

Contingency Tables Showing Predicted-Versus-Actual Groups Using Predictors Selected when the Predictand was a Single Predictand and Using Predictors Selected when the Predictand was a Pair of Predictands

Atlantic City 3-hr Visibility (Independent Sample)

		Actual					
	Group	1	2	3	4	5	Total
Using Predictors Selected When Predictand was Single	1	13	9	5	1	4	32
	2	2	5	4	0	9	20
	3	0	2	0	0	0	2
	4	0	0	3	4	0	7
	5	27	30	66	88	1772	1983
	Total	42	46	78	93	1785	2044

Hits = 1794 $\quad \bar{P} = .1800$

P_g	.0169	.0199	.0331	.0395	.0707

		Actual					
	Group	1	2	3	4	5	Total
Using Predictors Selected When Predictand was a Pair	1	10	7	4	0	3	24
	2	3	4	4	0	6	17
	3	0	3	0	0	0	3
	4	0	0	0	2	1	3
	5	29	32	70	91	1775	1997
	Total	42	46	78	93	1785	2044

Hits = 1791 $\quad \bar{P} = .1803$

P_g	.0167	.0199	.0334	.0395	.0708

EXPERIMENT 6. Effect of Restricting Selection to Predictors with Certain Minimum Occurrence Frequencies.

Purpose

This experiment was conducted to determine in a general way if there was any optimum limit to the number of occurrences to be permitted in a predictor before REEP would be allowed to include that predictor among those selected.

Procedure

Forecasts were made for 3-hr and 7-hr ceiling and visibility at Atlantic City Airport. Three different minimum number of occurrences of the predictors were specified in the separate runs, except for ceiling three hours in advance, which was tested for only two. The number of occurrences of the predictors specified were 10, 50, and 100. In the first case any predictor must have occurred at least ten times before it could be considered by REEP for selection, and similarly for the cases of 50 and 100. The forecasts made using the predictors selected by REEP in each case were then compared to determine the most favorable minimum occurrence frequency, comparing the $\overline{P}$ score and the number of correct forecasts in each case.

Data Sample

The developmental sample of 7956 observations is a random selection of hourly surface observations from the eight year period from 1 January 1949 to 31 December 1956. The verification sample of 2044 observations is a random selection from the two year period from 1 January 1957 to 31 December 1958. Observations are included in this sample for the 13 stations comprising this Atlantic City Airport network (Figure 1-1).

Predictands

Table 6-1 shows the five operationally significant categories of ceiling and visibility for the predictands used in this experiment. The specifications of the predictands are shown in Table 6-2.

Predictors

The 423 original dummy predictors used in this experiment are the same as those used in Experiment No. 1 (see Tables 1-3 and 1-4). The predictors with number of occurrences less than 100 are listed in Table 6-3. There are 90 such dummy predictors out of a total of 423. The number of observations in the developmental sample, as noted previously, is 7956.

Results

Contingency tables (Tables 6-4a to 6-4d) were formed to show the forecast-versus-observed frequencies for the 3-hr and 7-hr predictions of ceiling and visibility using the predictors selected by REEP when the conditions of 10, 50, and 100 occurrence minimum were applied to the selection procedure. The number of correct forecasts (hits) and the Brier and Allen $\bar{P}$ scores as well as the individual group contributions to the overall $\bar{P}$ score are given below each contingency table. The number of hits and the overall $\bar{P}$ scores are summarized in Table 6-5.

Conclusions

The results suggest that, when selection is restricted to predictors with a larger number of occurrences, the predictors are as good or better than those when restricted by fewer number of occurrences. In addition, there is no apparent deterioration in the ability to forecast low conditions by this restriction.

Table 6-1

Definition of Predictand Groups for Atlantic City Municipal Airport

Group Number	Ceiling Intervals
1	0 ft ≤ Ceiling < 200 ft
2	200 ft ≤ Ceiling < 500 ft
3	500 ft ≤ Ceiling < 1000 ft
4	1000 ft ≤ Ceiling < 3000 ft
5	3000 ft ≤ Ceiling

Group Number	Visibility Intervals
1	0 mi ≤ Visibility < 1/2 mi
2	1/2 mi ≤ Visibility < 1 mi
3	1 mi ≤ Visibility < 2 mi
4	2 mi ≤ Visibility < 3 mi
5	3 mi ≤ Visibility

Table 6-2

Specification of Predictands
Atlantic City Municipal Airport

	Predictand 1	2	3	4
Predictand variable	CIG	CIG	VIS	VIS
Forecast interval--hours (H)	3	7	3	7
Number of Groups (G)	5	5	5	5
Observations in Group 1 (n_1)	100	135	153	167
Observations in Group 2 (n_2)	281	272	125	130
Observations in Group 3 (n_3)	385	420	200	230
Observations in Group 4 (n_4)	660	619	286	283
Observations in Group 5 (n_5)	6530	6510	7192	7146
Total dependent sample size (N)	7956	7956	7956	7956
Number of independent observations in the dependent sample (n)				
Group 1	5500	5500	5500	5500
Group 2	5500	5500	5500	5500
Group 3	5500	5500	5500	5500
Group 4	5500	5500	5500	5500
Group 5	3000	3000	3000	3000
Total independent sample size (M)	2044	2044	2044	2044
Number of available predictors (P)	(423)	(423)	(423)	(423)
10 Occurrence Criteria	403	403	403	403
50 Occurrence Criteria	371	371	371	371
100 Occurrence Criteria	333	333	333	333

Table 6-3

Predictors in Developmental Sample (7956 Observations)
with Number of Occurrences Less Than 100

Predictors in dummy (zero-one) form

Predictor	Number of Occurrences
NEL WEA 5	46
NEL WEA 6	11
NEL WEA 7	43
NEL WEA 8	2
NEL WEA 11	5
PHL CIG 1	42
PHL VIS 1	68
PHL VIS 2	59
PHL WEA 4	41
PHL WEA 5	46
PHL WEA 6	7
PHL WEA 7	25
PHL WEA 8	5
PHL WEA 11	1
SBY CIG 1	71
SBY VIS 2	91
SBY WEA 5	36
SBY WEA 6	3
SBY WEA 7	60
SBY WEA 8	5
SBY WEA 11	1
MDT CIG 1	51
MDT CIG 2	90
MDT VIS 1	74
MDT WEA 4	72
MDT WEA 5	80
MDT WEA 6	17
MDT WEA 7	28
MDT WEA 8	13
MDT WEA 11	2

Predictor	Number of Occurrences
ACY CIG (1)	53
ACY CIG (2)	64
ACY CIG (3)	98
ACY CIG (4)	62
ACY VIS (1)	46
ACY VIS (2)	52
ACY VIS (3)	63
ACY VIS (5)	99
ACY WEA 5	42
ACY WEA 6	3
ACY WEA 7	36
ACY WEA 8	3
ACY WEA 11	5
DCA CIG 1	19
DCA CIG 2	81
DCA VIS 1	29
DCA VIS 2	16
DCA WEA 4	77
DCA WEA 5	45
DCA WEA 6	7
DCA WEA 7	50
DCA WEA 8	11
DCA WEA 11	0
ORF CIG 1	47
ORF VIS 1	67
ORF VIS 2	41
PSB WEA 7	67
PSB WEA 8	43
PSB WEA 11	45

Predictor	Number of Occurrences
BGM VIS 2	97
BGM WEA 4	84
BGM WEA 7	58
BGM WEA 8	37
BGM WEA 11	52
SWF CIG 1	81
SWF WEA 6	21
SWF WEA 7	56
SWF WEA 8	26
SWF WEA 11	8
RIC CIG 1	74
RIC VIS 1	91
RIC VIS 2	66
RIC WEA 5	38
RIC WEA 6	5
RIC WEA 7	52
RIC WEA 8	8
RIC WEA 11	0
IDL CIG 1	76
IDL VIS 1	90
IDL VIS 2	69
IDL WEA 5	64
IDL WEA 6	11
IDL WEA 7	37
IDL WEA 8	9
IDL WEA 11	6
DSM WEA 4	99
DSM WEA 6	40
DSM WEA 8	23
DSM WEA 10	70
DSM WEA 11	24

Table 6-4a

Contingency Tables Showing Predicted-Versus-Actual Groups When Selection of Predictors is Limited to Those With A Specified Minimum Occurrence Frequency

Atlantic City 3-hour Ceiling (Independent Sample)

REEP with Selection of Predictors Limited to 10-Occurrence Minimum

Group	Actual 1	2	3	4	5	Total
1	5	4	2	0	2	13
2	19	33	23	6	3	84
3	1	11	39	12	9	72
4	0	7	22	43	21	93
5	19	18	41	104	1600	1782
Total	44	73	127	165	1635	2044

Hits = 1720 $\quad \bar{P} = .2131$

P_g	.0165	.0278	.0433	.0592	.0663

REEP with Selection of Predictors Limited to 50-Occurrence Minimum

Group	Actual 1	2	3	4	5	Total
1						
2	(Prediction not made for					
3						
4	3-hr Ceiling)					
5						
Total						

Hits = $\quad \bar{P} =$

P_g

REEP with Selection of Predictors Limited to 100-Occurrence Minimum

Group	Actual 1	2	3	4	5	Total
1	10	5	2	0	3	20
2	18	30	25	6	2	81
3	0	13	41	12	8	74
4	0	7	20	46	18	91
5	16	18	39	101	1604	1778
Total	44	73	127	165	1635	2044

Hits = 1731 $\quad \bar{P} = .2110$

P_g	.0161	.0272	.0429	.0592	.0656

Table 6-4b

Contingency Tables Showing Predicted-Versus-Actual Groups When Selection of Predictors is Limited to Those With A Specified Minimum Occurrence Frequency

Atlantic City 7-hour Ceiling (Independent Sample)

REEP with Selection of Predictors Limited to 10-Occurrence Minimum

	Actual					
Group	1	2	3	4	5	Total
1	0	0	0	0	0	0
2	9	17	7	3	2	38
3	3	15	32	10	6	66
4	5	8	19	16	7	55
5	24	39	89	129	1604	1885
Total	41	79	147	158	1619	2044

Hits = 1669 $\quad \bar{P} = .2567$

P_g	.0177	.0325	.0539	.0658	.0868

REEP with Selection of Predictors Limited to 50-Occurrence Minimum

	Actual					
Group	1	2	3	4	5	Total
1	0	0	0	0	0	0
2	7	17	8	3	3	38
3	6	16	34	11	7	74
4	5	6	18	13	8	50
5	23	40	87	131	1601	1882
Total	41	79	147	158	1619	2044

Hits = 1665 $\quad \bar{P} = .2567$

P_g	.0177	.0325	.0540	.0660	.0866

REEP with Selection of Predictors Limited to 100-Occurrence Minimum

	Actual					
Group	1	2	3	4	5	Total
1	0	0	0	0	0	0
2	9	16	10	5	2	42
3	5	16	33	10	6	70
4	5	5	19	16	6	51
5	22	42	85	127	1605	1881
Total	41	79	147	158	1619	2044

Hits = 1670 $\quad \bar{P} = .2572$

P_g	.0178	.0326	.0538	.0662	.0868

Table 6-4c

Contingency Tables Showing Predicted-Versus-Actual Groups When Selection of Predictors is Limited to Those With A Specified Minimum Occurrence Frequency

Atlantic City 3-hour Visibility (Independent Sample)

REEP with Selection of Predictors Limited to 10-Occurrence Minimum

	Actual					
Group	1	2	3	4	5	Total
1	11	6	3	0	5	25
2	3	6	7	0	5	21
3	1	3	0	0	3	7
4	0	0	0	6	1	7
5	27	31	68	87	1771	1984
Total	42	46	78	93	1785	2044

Hits = 1794 $\quad \bar{P} = .1799$

P_g	.0166	.0197	.0328	.0393	.0715

REEP with Selection of Predictors Limited to 50-Occurrence Minimum

	Actual					
Group	1	2	3	4	5	Total
1	11	8	5	0	3	27
2	1	4	5	0	6	16
3	1	3	0	0	1	5
4	1	0	1	6	1	9
5	28	31	67	87	1774	1987
Total	42	46	78	93	1785	2044

Hits = 1795 $\quad \bar{P} = .1793$

P_g	.0168	.0199	.0327	.0392	.0707

REEP with Selection of Predictors Limited to 100-Occurrence Minimum

	Actual					
Group	1	2	3	4	5	Total
1	19	11	7	0	6	43
2	1	4	1	0	7	13
3	0	1	0	0	0	1
4	1	0	0	6	0	7
5	21	30	70	87	1772	1980
Total	42	46	78	93	1785	2044

Hits = 1801 $\quad \bar{P} = .1799$

P_g	.0169	.0201	.0332	.0393	.0705

Table 6-4d

Contingency Tables Showing Predicted-Versus-Actual Groups When Selection of Predictors is Limited to Those With A Specified Minimum Occurrence Frequency

Atlantic City 7-hour Visibility (Independent Sample)

REEP with Selection of Predictors Limited to 10-Occurrence Minimum

	Actual					
Group	1	2	3	4	5	Total
1	0	0	0	1	1	2
2	1	0	0	0	0	1
3	1	0	0	0	0	1
4	0	0	0	0	0	0
5	41	38	90	98	1773	2040
Total	43	38	90	99	1774	2044

Hits = 1773 $\quad \bar{P} = .2161$

P_g	.0195	.0181	.0404	.0447	.0934

REEP with Selection of Predictors Limited to 50-Occurrence Minimum

	Actual					
Group	1	2	3	4	5	Total
1	0	0	0	1	1	2
2	1	0	0	0	0	1
3	1	0	0	0	0	1
4	0	0	0	0	0	0
5	41	38	90	98	1773	2040
Total	43	38	90	99	1774	2044

Hits = 1773 $\quad \bar{P} = .2162$

P_g	.0195	.0180	.0404	.0447	.0935

REEP with Selection of Predictors Limited to 100-Occurrence Minimum

	Actual					
Group	1	2	3	4	5	Total
1	0	0	1	0	1	2
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	43	38	89	99	1773	2042
Total	43	38	90	99	1774	2044

Hits = 1773 $\quad \bar{P} = .2144$

P_g	.0193	.0180	.0401	.0447	.0923

Table 6-5

Comparison Between Predictions Made When Selection of Predictors is Limited to a Minimum Occurrence Frequency of 10, 50, and 100.

ACY Ceiling - 3-hr Forecast

Hits	10 Occurrence Minimum	1720
	50 Occurrence Minimum	--
	100 Occurrence Minimum	1731*
$\bar{P}$ Score	10 Occurrence Minimum	.2131
	50 Occurrence Minimum	--
	100 Occurrence Minimum	.2110*

ACY Ceiling - 7-hr Forecast

Hits	10 Occurrence Minimum	1669
	50 Occurrence Minimum	1665
	100 Occurrence Minimum	1670*
$\bar{P}$ Score	10 Occurrence Minimum	.2567*
	50 Occurrence Minimum	.2567*
	100 Occurrence Minimum	.2572

ACY Visibility - 3-hr Forecast

Hits	10 Occurrence Minimum	1794
	50 Occurrence Minimum	1795
	100 Occurrence Minimum	1801*
$\bar{P}$ Score	10 Occurrence Minimum	.1799
	50 Occurrence Minimum	.1793*
	100 Occurrence Minimum	.1799

ACY Visibility - 7-hr Forecast

Hits	10 Occurrence Minimum	1773
	50 Occurrence Minimum	1773
	100 Occurrence Minimum	1773
$\bar{P}$ Score	10 Occurrence Minimum	.2161
	50 Occurrence Minimum	.2162
	100 Occurrence Minimum	.2144*

Asterisks are used to denote superiority.

EXPERIMENT 7. Single-Station Predictors: Raw Dummies and Boolean Combinations.

Purpose

To investigate the accuracy of forecasts using only single-station predictors. In the event the accuracy is sufficiently competitive, a generalized operator might be developed where spacial analyses of network parameters would not be required.

Procedure

From the 423 predictors used in the Atlantic City network of 13 stations (see Table 1-3), the 55 predictors for only Atlantic City predictors (not including TOD and DOY) were extracted. Selections and predictions were first made for ACY ceiling and visibility for three and seven hours using the 55 dummy predictors alone.

Next, those dummy predictors with 400 or more occurrences were then formed into all possible combinations of pairs joined by a logical "AND" including the combination of each predictor with itself. An exception to this procedure was made in figuring the number of occurrences under the Weather Element heading. In this case the weather dummies from 2 to 8 (precipitation) were grouped together to represent a precipitation predictor. Combining the dummy predictors in this way resulted in the formation of 467 new predictors. Selections and predictions were again made for ACY ceiling and visibility at three and seven hours in advance using the 467 Boolean-paired predictors. The results were compared with those from the prediction runs with the raw dummy predictors, as well as with the results using all the network predictors (Experiment No. 1).

Predictands

The predictands, as used in the earlier experiments, are subdivided into five operationally significant intervals of ceiling and visibility as shown in Table 7-1. The specifications of the predictands are shown in Tables 7-2a and 7-2b.

Predictors

As noted previously the predictors used in this experiment were extracted from the 423 predictors used in the Atlantic City network of 13 stations. The 55 dummy predictors used in the first prediction runs are shown in Table 7-3 together with the number of occurrences of each in the developmental sample of 7956 observations. Table 7-3 also lists the specifications and limits of the dummy variables. Those predictors with 400 occurrences or more, plus the time of day (TOD) and day of year (DOY), were formed into all possible pairs joined by a logical "AND" resulting in 467 Boolean-paired predictors.

Data Sample

As in the earlier experiments the 10,000 observations used in this experiment were randomly selected from the ten year period 1 January 1949 to 31 December 1958. The developmental sample has 7956 observations from the first eight years of the ten year period, and the verification sample has 2044 observations from the last two years.

Results

The four sets of predictors selected by REEP, using the 55 predictors for ACY are listed in Tables 7-4a to 7-4d. The predictors selected using the 467 Boolean-paired predictors are shown in Tables 7-5a to 7-5d. Also shown in each of these tables are coefficients in the REEP

regression equation for each group of the predictand. Contingency tables (Tables 7-6a to 7-6d) show the forecast-versus-observed frequencies for the 3-hr and 7-hr predictions of ceiling and visibility using the 55 dummy predictors for ACY listed in Table 7-3 and also using the 467 Boolean combination of predictors previously described. The number of correct forecasts (hits) and the Brier and Allen $\overline{P}$ scores as well as the individual group contributions to the overall $\overline{P}$ score are given below each contingency table. Tables 7-7a and 7-7b summarize the hits and $\overline{P}$ scores for this experiment. In addition, these tables show how the results of this experiment compare with those of Experiment No. 1, in which all the dummy predictors for the 13 station Atlantic City network were used, and with the results of a straight persistence forecast.

Conclusions

1) The single-station raw dummy predictors give results which are better than persistence.

2) The single-station Boolean predictors are as good or better than the raw predictors except for the $\overline{P}$ score in predicting 3-hr ceiling. Low conditions are forecast more often and more accurately with Boolean predictors.

3) Although the selected network predictors are quite a bit better than either single-station set of predictors, the Boolean set is consistently better than the network in predicting visibility.

4) The results are sufficiently encouraging to suggest that more work be done. In addition, it seems worth attempting to combine the set of network and Boolean predictors in a separate experiment (see Experiment No. 8).

Table 7-1

Definition of Predictand Groups for Atlantic City Municipal Airport

Group Number	Ceiling Intervals
1	0 ft $\leq$ Ceiling $<$ 200 ft
2	200 ft $\leq$ Ceiling $<$ 500 ft
3	500 ft $\leq$ Ceiling $<$ 1000 ft
4	1000 ft $\leq$ Ceiling $<$ 3000 ft
5	3000 ft $\leq$ Ceiling

Group Number	Visibility Intervals
1	0 mi $\leq$ Visibility $<$ 1/2 mi
2	1/2 mi $\leq$ Visibility $<$ 1 mi
3	1 mi $\leq$ Visibility $<$ 2 mi
4	2 mi $\leq$ Visibility $<$ 3 mi
5	3 mi $\leq$ Visibility

Table 7-2a

Specification of Predictands - Atlantic City Municipal Airport

(As used in prediction runs using raw predictors)

	Predictand 1	2	3	4
Predictand variable	CIG	CIG	VIS	VIS
Forecast interval--hours (H)	3	7	3	7
Number of Groups (G)	5	5	5	5
Observations in Group 1 (n_1)	100	135	153	167
Observations in Group 2 (n_2)	281	272	125	130
Observations in Group 3 (n_3)	385	420	200	230
Observations in Group 4 (n_4)	660	619	286	283
Observations in Group 5 (n_5)	6530	6510	7192	7146
Total dependent sample size (N)	7956	7956	7956	7956
Number of independent observations in the dependent sample (n)	7956	7956	7956	7956
Total independent sample size (M)	2044	2044	2044	2044
Number of available predictors (P)	55	55	55	55

Table 7-2b

Specification of Predictands - Atlantic City Municipal Airport

(As used in prediction runs using Boolean predictors)

	Predictand 1	2	3	4
Predictand variable	CIG	CIG	VIS	VIS
Forecast interval--hours (H)	3	7	3	7
Number of Groups (G)	5	5	5	5
Observations in Group 1 (n_1)	100	135	153	167
Observations in Group 2 (n_2)	281	272	125	130
Observations in Group 3 (n_3)	385	420	200	230
Observations in Group 4 (n_4)	660	619	286	283
Observations in Group 5 (n_5)	6530	6510	7192	7146
Total dependent sample size (N)	7956	7956	7956	7956
Number of independent observations in the dependent sample (n)	2652	2652	2652	2652
Total independent sample size (M)	2044	2044	2044	2044
Number of available predictors (P)	467	467	467	467

Table 7-2b

Specification of Predictands - Atlantic City Municipal Airport

(As used in prediction runs using Boolean predictors)

	Predictand 1	2	3	4
Predictand variable	CIG	CIG	VIS	VIS
Forecast interval--hours (H)	3	7	3	7
Number of Groups (G)	5	5	5	5
Observations in Group 1 (n_1)	100	135	153	167
Observations in Group 2 (n_2)	281	272	125	130
Observations in Group 3 (n_3)	385	420	200	230
Observations in Group 4 (n_4)	660	619	286	283
Observations in Group 5 (n_5)	6530	6510	7192	7146
Total dependent sample size (N)	7956	7956	7956	7956
Number of independent observations in the dependent sample (n)	2652	2652	2652	2652
Total independent sample size (M)	2044	2044	2044	2044
Number of available predictors (P)	467	467	467	467

Table 7-3

Predictors Available for Single-Station Predictions
Using Dummy Predictors for the Forecast Station (ACY)

Element	No. of Dummies	Dummy No.	Specifications and Limits	Number of Occurrences
			Ceiling Height*, ft:	
CIG	5	1.	$0 \leq CIG < 200$	109
		2.	$200 \leq CIG < 500$	263
		3.	$500 \leq CIG < 1000$	410
		4.	$1000 \leq CIG < 3000$	661
		5.	$3000 \leq CIG$	6513
	10	(1)	$CIG = 100$	53
		(2)	$CIG = 200$	64
		(3)	$CIG = 400$	98
		(4)	$CIG = 500$	108
		(5)	$CIG = 900$	62
		(6)	$1500 \leq CIG < 2000$	146
		(7)	$2500 \leq CIG < 3000$	123
		(8)	$3000 \leq CIG < 5000$	596
		(9)	$5000 \leq CIG < 10,000$	722
		(10)	CIG = Unlimited	4039
			Visibility*, mi:	
VIS	5	1.	$0 \leq VIS < 1/2$	139
		2.	$1/2 \leq VIS < 1$	115
		3.	$1 \leq VIS < 2$	213
		4.	$2 \leq VIS < 3$	274
		5.	$3 \leq VIS$	7215
	10	(1)	$1/4 \leq VIS < 1/2$	46
		(2)	$1/2 \leq VIS < 3/4$	52
		(3)	$3/4 \leq VIS < 1$	63
		(4)	$1 \leq VIS < 1-1/2$	114
		(5)	$1-1/2 \leq VIS < 2$	99

*Parentheses ( ) used to help distinguish the two forms of dummies.

Table 7-3 - Continued

Element	No. of Dummies	Dummy No.	Specifications and Limits	Number of Occurrences
			Visibility*. mi:	
		(6)	$2 \leq$ VIS $<$ 3	274
		(7)	$3 \leq$ VIS $<$ 4	334
		(8)	$4 \leq$ VIS $<$ 5	282
		(9)	$5 \leq$ VIS $<$ 6	370
		(10)	$15 \leq$ VIS	1575
			Wind:	
WND	5	1.	Calm to 3 Knots, any direction	723
		2.	23° to 67°, 3 Knots < Speed	854
		3.	68° to 202°, 3 Knots < Speed	1727
		4.	203° to 247°, 3 Knots < Speed	1656
		5.	248° to 22°, 3 Knots < Speed	2976
			Relative Humidity:	
RLH	3	1.	75 per cent to 85 per cent	1387
		2.	86 per cent to 93 per cent	1418
		3.	94 per cent to 100 per cent	1262
			Total Cloud Amount:	
TCA	2	1.	5/10 to 9/10 (incl.)	2026
		2.	More than 9/10	2633
			Weather Element**:	
WEA	12	1.	No Weather	5721
		2.	Rain	395
		3.	Rain Showers	144
		4.	Drizzle	106
		5.	Snow	42
		6.	Snow Showers	3
		7.	Thunderstorm, Hail	36

*Parentheses ( ) used to help distinguish the two forms of dummies.

**See Table 1-4 for a complete list of Weather Elements.

Table 7-3 - Continued

Element	No. of Dummies	Dummy No.	Specifications and Limits	Number of Occurrences
			Weather Elements**:	
		8.	Freezing Rain, Sleet	3
		9.	Fog	689
		10.	Ground Fog	439
		11.	Blowing Snow, Blowing Sand	5
		12.	Smoke, Haze, Dust	816
			Sky Condition, Lower:	
SCL	3	1.	Sky obscured	209
		2.	Thin broken clouds Thin overcast	380
		3.	Scattered clouds Dark scattered clouds	2267

**See Table 1-4 for a complete list of Weather Elements.

Table 7-4a

Selected Predictors and REEP Equations Using
Single-Station Predictors in Raw Dummy Form

Atlantic City 3-hr Ceiling

Selected Predictors				Coefficients of the REEP Equations				
Order	Station	Element	Dummy	A_1	A_2	A_3	A_4	A_5
1	ACY	CIG	5	-.001	-.020	-.106	-.360	.487
2	ACY	CIG	1	.365	.278	.022	-.477	-.188
3	ACY	CIG	2	.035	.331	.017	-.301	-.082
4	ACY	CIG	3	-.006	.090	.175	-.208	-.052
5	ACY	TCA	2	.002	-.005	.023	.050	-.070
6	ACY	WEA	2	-.025	.083	.063	.018	-.140
7	ACY	VIS	5	-.061	.026	.041	-.040	.035
8	ACY	WND	5	-.006	-.026	-.032	-.022	.086
9	ACY	CIG	(1)	-.187	.050	-.106	.079	.164
10	ACY	CIG	(7)	.004	.012	-.086	-.129	.200
11	ACY	WND	4	-.007	-.028	-.033	-.013	.081
12	ACY	VIS	(1)	.099	-.024	.001	.040	-.116
13	ACY	WEA	9	.004	.055	-.001	.001	-.059
14	ACY	VIS	(4)	-.055	-.025	.161	-.049	-.032
15	ACY	WEA	7	-.014	-.035	-.077	-.180	.306
16	ACY	CIG	(8)	.001	-.002	.014	.082	-.095
17	ACY	VIS	4	-.038	.018	.028	-.026	.018
18	ACY	RLH	2	.005	.013	.026	.038	-.082
19	ACY	CIG	(6)	-.008	-.016	-.076	.142	-.042
20	ACY	WND	1	.006	-.021	-.042	-.014	.071
21	ACY	RLH	3	.013	.028	.038	.002	-.082
22	ACY	RLH	1	-.001	-.001	.011	.039	-.049
23	ACY	CIG	(4)	.004	-.004	.109	-.109	-.001
24	ACY	CIG	(5)	.001	-.100	-.018	.041	.075
25	ACY	WEA	4	-.006	.009	.082	.028	-.112
26	ACY	SCL	3	.000	.001	.008	.019	-.028
27	ACY	VIS	2	.035	.022	.097	-.066	-.088
28	ACY	SCL	1	.002	-.073	-.036	.042	.065
29	ACY	CIG	(2)	.012	.084	-.054	-.042	-.000
30	ACY	CIG	(9)	-.001	.001	-.002	.035	-.032
		Additive Constant:		.066	.014	.083	.409	.428

Table 7-4b

Selected Predictors and REEP Equations Using
Single-Station Predictors in Raw Dummy Form

Atlantic City 7-hr Ceiling

Selected Predictors				Coefficients of the REEP Equations				
Order	Station	Element	Dummy	A_1	A_2	A_3	A_4	A_5
1	ACY	CIG	5	-.010	-.110	-.137	-.079	.336
2	ACY	TCA	2	.000	.011	.037	.060	-.108
3	ACY	WND	5	-.025	-.047	-.052	-.039	.163
4	ACY	WND	4	-.022	-.044	-.053	-.035	.155
5	ACY	RLH	3	.010	-.000	.046	.086	-.142
6	ACY	CIG	2	.006	.090	-.016	-.090	.011
7	ACY	SCL	1	.047	-.074	-.057	-.019	.103
8	ACY	RLH	2	.009	.018	.033	.050	-.110
9	ACY	WEA	2	.001	.079	.075	-.037	-.117
10	ACY	WND	1	-.011	-.032	-.047	-.021	.111
11	ACY	CIG	4	-.008	-.048	-.088	.060	.085
12	ACY	RLH	1	.007	.008	.019	.032	-.066
13	ACY	WEA	9	.025	.004	.010	-.000	-.039
14	ACY	VIS	(3)	.054	.008	.090	-.135	-.017
15	ACY	WEA	4	.046	.044	.070	-.016	-.143
16	ACY	CIG	(8)	.004	.005	.012	.039	-.060
17	ACY	WEA	7	.010	.002	-.130	-.109	.228
18	ACY	CIG	1	.127	.035	-.004	.025	-.184
19	ACY	CIG	(1)	-.122	-.010	.042	-.094	.184
20	ACY	VIS	2	.039	.047	-.099	.119	-.106
21	ACY	VIS	4	.029	.014	.007	-.045	-.005
22	ACY	CIG	(5)	-.048	-.079	-.033	.116	.044
23	ACY	CIG	(7)	.026	-.073	-.005	-.034	.086
24	ACY	CIG	(6)	.011	-.063	-.020	.070	.002
25	ACY	WEA	3	-.001	-.020	.063	.019	-.061
		Additive Constant:		.029	.146	.174	.115	.536

Table 7-4c

Selected Predictors and REEP Equations Using
Single-Station Predictors in Raw Dummy Form

Atlantic City 3-hr Visibility

Selected Predictors				Coefficients of the REEP Equations				
Order	Station	Element	Dummy	A_1	A_2	A_3	A_4	A_5
1	ACY	VIS	5	-.025	-.014	-.067	-.098	.204
2	ACY	VIS	1	.285	.041	-.087	-.150	-.089
3	ACY	RLH	3	.034	.026	.040	.035	-.134
4	ACY	VIS	2	.156	.189	-.016	-.084	-.245
5	ACY	WEA	1	-.036	.002	.006	-.040	.069
6	ACY	WEA	9	.000	.018	.042	.047	-.106
7	ACY	RLH	2	.011	.006	.018	.027	-.062
8	ACY	CIG	1	.217	.017	.119	.041	-.395
9	ACY	CIG	(1)	-.138	-.056	-.086	-.065	.344
10	ACY	WEA	2	-.062	-.013	.032	.026	.017
11	ACY	WEA	5	-.060	.088	.154	-.077	-.135
12	ACY	CIG	(2)	-.027	-.007	.119	.036	-.122
13	ACY	VIS	3	.053	.041	-.053	-.005	-.036
14	ACY	SCL	1	-.071	.030	-.031	-.026	.098
15	ACY	VIS	(1)	-.052	.022	.140	-.062	-.048
16	ACY	VIS	(8)	-.012	.017	.005	.055	-.064
17	ACY	WEA	4	-.056	.025	.067	.006	-.044
18	ACY	WEA	7	-.081	-.010	-.058	-.036	.184
19	ACY	WND	5	-.005	-.010	-.007	-.014	.038
20	ACY	WND	4	-.004	-.010	-.016	-.005	.035
21	ACY	VIS	(4)	.030	-.017	.082	-.103	.008
22	ACY	WEA	12	-.029	.006	.018	-.014	.019
23	ACY	CIG	3	-.029	-.012	.011	.018	.012
		Additive Constant:		.064	.021	.069	.149	.697

Table 7-4d

Selected Predictors and REEP Equations Using
Single-Station Predictors in Raw Dummy Form

Atlantic City 7-hr Visibility

Selected Predictors				Coefficients of the REEP Equations				
Order	Station	Element	Dummy	A_1	A_2	A_3	A_4	A_5
1	ACY	WEA	9	.012	.001	.044	.009	-.066
2	ACY	WEA	1	-.011	-.002	-.011	-.028	.052
3	ACY	WND	5	-.013	-.020	-.018	-.027	.078
4	ACY	WND	4	-.003	-.017	-.015	-.020	.055
5	ACY	CIG	2	.014	.037	.053	.011	-.115
6	ACY	VIS	5	-.048	-.011	.002	.011	.046
7	ACY	RLH	2	.014	.012	.022	.021	-.069
8	ACY	CIG	3	.005	.004	.046	.007	-.062
9	ACY	VIS	2	-.029	.092	-.041	-.010	-.012
10	ACY	TCA	2	-.009	.001	.014	.022	-.028
11	ACY	CIG	1	.121	-.070	.072	.063	-.186
12	ACY	CIG	(1)	-.128	-.023	-.040	.065	.126
13	ACY	VIS	(3)	.129	-.020	.080	-.036	-.153
14	ACY	SCL	1	.019	.068	-.054	-.032	-.002
15	ACY	WND	3	.021	.000	.000	-.003	-.018
16	ACY	WEA	4	.005	.051	-.006	.008	-.058
17	ACY	RLH	3	.015	.005	.010	.010	-.040
18	ACY	WEA	5	-.055	.046	.034	-.010	-.014
19	ACY	WEA	2	-.016	.000	.030	.038	-.052
		Additive Constant:		.069	.031	.026	.045	.829

Table 7-5a

Selected Predictors and REEP Equations Using
Single-Station Predictors in Boolean Pairs

Atlantic City 3-hr Ceiling

Order	Selected Predictors Station	Element	Dummy	Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1	ACY	CIG	5	.063	-.029	-.064	-.287	.318
2	ACY	CIG	1	.414	.215	.009	-.378	-.260
3	ACY	CIG	2	.068	.304	.039	-.260	-.151
4	ACY	CIG	3	.005	.059	.221	-.209	-.076
5	ACY and ACY	TOD 1930E-0130E WEA 9		.171	.010	-.103	-.082	.004
6	ACY	TCA	2	-.005	-.021	-.010	.063	-.027
7	ACY and ACY	WEA 2-8 WEA 9		-.039	.108	.045	-.046	-.068
8	ACY and ACY	WND 2 TCA 2		.014	.059	.085	.038	-.196
9	ACY and ACY	WND 3 TCA 2		.015	.055	.069	-.007	-.132
10	ACY and ACY	CIG 5 VIS 5		-.067	-.021	-.048	-.048	.184
11	ACY	CIG	(1)	-.197	.040	-.058	.075	.139
12	ACY	WEA	2	-.010	.079	.039	.028	-.136
13	ACY	CIG	(7)	.006	.010	-.073	-.171	.227
14	ACY and ACY	CIG (8) TCA 2		.001	.004	.024	.103	-.132
15	ACY and ACY	TOD 0730E-1330E CIG 4		-.007	-.061	-.035	.150	-.047
16	ACY and ACY	WEA 2-8 DOY 151-300		-.004	-.090	.076	-.016	.035
		Additive Constant:		.005	.054	.116	.353	.472

Table 7-5b

Selected Predictors and REEP Equations Using
Single-Station Predictors in Boolean Pairs

Atlantic City 7-hr Ceiling

Order	Selected Predictors				Coefficients of the REEP Equations $-A_1$	A_2	A_3	A_4	A_5
1		ACY	CIG	5	-.005	-.049	-.069	-.129	.252
2		ACY	TCA	2	.025	.076	.123	.113	-.336
3	and	ACY	TOD	1930E-0130E	.175	.024	-.031	-.095	-.074
		ACY	WEA	9					
4	and	ACY	WND	5	-.037	-.104	-.109	-.063	.313
		ACY	TCA	2					
5	and	ACY	WND	4	-.034	-.092	-.115	-.050	.292
		ACY	TCA	2					
6		ACY	RLH	3	.023	.054	.006	.045	-.128
7		ACY	CIG	2	-.006	.105	.007	-.089	-.016
8		ACY	RLH	2	.009	.023	.034	.030	-.095
9		ACY	WEA	2	-.005	.064	.044	-.018	-.085
10	and	ACY	TOD	1330E-1930E	.104	-.002	-.077	-.053	.027
		ACY	WEA	9					
11	and	ACY	WND	1	-.008	-.063	-.094	-.018	.183
		ACY	TCA	2					
12	and	ACY	WND	3	.020	.014	.028	.050	-.113
		ACY	TCA	1					
13	and	ACY	DOY	151-300	-.015	-.069	.100	.021	-.037
		ACY	RLH	3					
14	and	ACY	WEA	2-8	-.007	.025	.132	-.059	-.091
		ACY	WEA	9					
15		ACY	SCL	1	.072	-.026	-.012	-.026	-.007
16	and	ACY	TOD	0130E-0730E	-.015	-.015	-.023	.071	-.017
		ACY	DOY	151-300					
			Additive Constant		.008	.051	.069	.140	.732

Table 7-5c

Selected Predictors and REEP Equations Using Single-Station Predictors in Boolean Pairs

Atlantic City 3-hr Visibility

Order	Selected Predictors Station	Element	Dummy	Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1	ACY	VIS	5	-.039	-.029	-.076	-.082	.227
2	ACY	VIS	1	.290	.042	-.008	-.134	-.190
3	ACY	RLH	3	.042	.036	.054	.056	-.187
4	ACY	VIS	2	.114	.179	-.007	-.062	-.224
5	ACY and ACY	CIG WEA	5 1	-.002	-.008	-.013	-.042	.064
6	ACY and ACY	TOD WEA	1930E-0130E 9	.133	.064	.044	-.006	-.236
7	ACY and ACY	TOD WEA	1930E-0130E 10	.121	-.008	.012	-.038	-.087
8	ACY	RLH	2	.007	.005	.019	.033	-.065
9	ACY and ACY	WND RLH	4 3	-.013	-.037	-.081	-.016	.147
10	ACY and ACY	WEA RLH	2-8 2	-.068	-.017	.028	.008	.048
11	ACY and ACY	WEA TOD	2-8 0130E-0730E	.011	.003	.053	.087	-.153
		Additive Constant:		.041	.038	.090	.126	.706

Table 7-5d

Selected Predictors and REEP Equations Using
Single-Station Predictors in Boolean Pairs

Atlantic City 7-hr Visibility

Order	Selected Predictors					Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1		ACY	CIG	5		-.027	-.022	-.057	-.043	.148
	and	ACY	VIS	5						
2		ACY	TOD	1930E-0130E		.092	.021	.053	.040	-.205
	and	ACY	RLH	3						
3		ACY	TCA	2		.059	.044	.059	.040	-.201
	and	ACY	WEA	9						
4		ACY	TOD	1930E-0130E		.026	.020	.045	.052	-.143
	and	ACY	RLH	2						
5		ACY	TOD	1330E-1930E		.043	.028	.037	.042	-.150
	and	ACY	WEA	12						
6		ACY	WND	5		-.018	-.019	-.018	-.028	.083
7		ACY	WND	4		-.013	-.017	-.017	-.020	.067
8		ACY	VIS	2		.067	.095	-.030	-.042	-.090
9		ACY	TOD	0130E-0730E		-.068	-.055	-.028	-.017	.167
	and	ACY	WEA	9						
10		ACY	CIG	4		-.026	-.020	-.039	-.026	.111
	and	ACY	VIS	5						
					Additive Constant:	.042	.040	.077	.077	.765

Table 7-6a

Contingency Tables Showing Predicted-Versus-Actual Groups When Single Station Predictors are in Raw Dummy Form and When Single Station Predictors are in Boolean Form

Atlantic City 3-hour Ceiling (Independent Sample)

		Actual					
	Group	1	2	3	4	5	Total
REEP Single Station Predictors in Raw Dummy Form	1	9	4	2	0	3	18
	2	20	36	27	10	10	103
	3	1	11	45	14	12	83
	4	1	9	20	47	28	105
	5	13	13	33	94	1582	1735
	Total	44	73	127	165	1635	2044

Hits = 1719 $\quad \bar{P} = .2246$

P_g	.0163	.0275	.0440	.0613	.0755

		Actual					
	Group	1	2	3	4	5	Total
REEP Single Station Predictors in Boolean Form (Pairs)	1	12	3	1	0	1	17
	2	19	38	27	10	10	104
	3	0	9	42	10	6	67
	4	1	7	16	40	29	93
	5	12	16	41	105	1589	1763
	Total	44	73	127	165	1635	2044

Hits = 1721 $\quad \bar{P} = .2275$

P_g	.0158	.0277	.0442	.0626	.0772

Table 7-6b

Contingency Tables Showing Predicted-Versus-Actual Groups When Single Station Predictors are in Raw Dummy Form and When Single Station Predictors are in Boolean Form

Atlantic City 7-hour Ceiling (Independent Sample)

REEP Single Station Predictors in Raw Dummy Form

Group	Actual 1	2	3	4	5	Total
1	0	0	1	1	0	2
2	3	9	9	5	4	30
3	8	11	24	12	8	63
4	2	5	5	3	4	19
5	28	54	108	137	1603	1930
Total	41	79	147	158	1619	2044

Hits = 1639 $\quad \bar{P} = .2678$

P_g	.0187	.0336	.0564	.0658	.0932

REEP Single Station Predictors in Boolean Form (Pairs)

Group	Actual 1	2	3	4	5	Total
1	2	1	0	2	3	8
2	6	11	9	5	3	34
3	6	12	26	9	5	58
4	4	7	13	8	6	38
5	23	48	99	134	1602	1906
Total	41	79	147	158	1619	2044

Hits = 1649 $\quad \bar{P} = .2676$

P_g	.0184	.0326	.0567	.0660	.0939

Table 7-6c

Contingency Tables Showing Predicted-Versus-Actual Groups When Single Station Predictors are in Raw Dummy Form and When Single Station Predictors are in Boolean Form

Atlantic City 3-hour Visibility (Independent Sample)

REEP Single Station Predictors in Raw Dummy Form

	Actual					
Group	1	2	3	4	5	Total
1	8	6	3	0	3	20
2	0	2	6	0	5	13
3	1	3	0	0	0	4
4	0	0	0	0	0	0
5	33	35	69	93	1777	2007
Total	42	46	78	93	1785	2044

Hits = 1787 $\quad \bar{P} = .1796$

P_g	.0167	.0197	.0326	.0392	.0713

REEP Single Station Predictors in Boolean Form (Pairs)

	Actual					
Group	1	2	3	4	5	Total
1	16	9	7	1	4	37
2	0	1	0	0	3	4
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	26	36	71	92	1778	2003
Total	42	46	78	93	1785	2044

Hits = 1795 $\quad \bar{P} = .1788$

P_g	.0165	.0201	.0327	.0389	.0705

Table 7-6d

Contingency Tables Showing Predicted-Versus-Actual Groups When Single Station Predictors are in Raw Dummy Form and When Single Station Predictors are in Boolean Form

Atlantic City 7-hour Visibility (Independent Sample)

REEP Single Station Predictors in Raw Dummy Form

		Actual					
	Group	1	2	3	4	5	Total
	1	0	0	0	0	0	0
	2	0	0	0	0	0	0
	3	0	0	0	0	0	0
	4	0	0	0	0	0	0
	5	43	38	90	99	1774	2044
	Total	43	38	90	99	1774	2044

Hits = 1774 $\quad \bar{P} = .2170$

P_g	.0200	.0180	.0401	.0444	.0945

REEP Single Station Predictors in Boolean Form (Pairs)

		Actual					
	Group	1	2	3	4	5	Total
	1	0	0	0	0	0	0
	2	0	0	0	0	0	0
	3	0	0	0	0	0	0
	4	0	0	0	0	0	0
	5	43	38	90	99	1774	2044
	Total	43	38	90	99	1774	2044

Hits = 1774 $\quad \bar{P} = .2140$

P_g	.0196	.0178	.0398	.0443	.0924

Table 7-7a

Comparison Between REEP Forecasts Using Single-Station Predictors in Raw Dummy Form, Single-Station Predictors in Boolean Pairs, All the Dummy Predictors from the 13-Station Network*, and Persistence*.

ACY Ceiling, 3-hr Forecast (Independent Sample)

Hits	Single-Station Dummies	1719
	Single-Station Boolean Pairs	1721
	Network Dummies	1728**
	Persistence	1702
$\overline{P}$ Score	Single-Station Dummies	.2246
	Single-Station Boolean Pairs	.2275
	Network Dummies	.2135**
	Persistence	.2385

ACY Ceiling, 7-hr Forecast (Independent Sample)

Hits	Single-Station Dummies	1639
	Single-Station Boolean Pairs	1649
	Network Dummies	1665**
	Persistence	1588
$\overline{P}$ Score	Single-Station Dummies	.2678
	Single-Station Boolean Pairs	.2676
	Network Dummies	.2600**
	Persistence	.2899

*See Experiment No. 1.

**Asterisks are used to denote superiority.

Table 7-7b

Comparison Between REEP Forecasts Using Single-Station Predictors in Raw Dummy Form, Single-Station Predictors in Boolean Pairs, All the Dummy Predictors from the 13-Station Network*, and Persistence*.

ACY Visibility, 3-hr Forecast (Independent Sample)

Hits	Single-Station Dummies	1787
	Single-Station Boolean Pairs	1795**
	Network Dummies	1794
	Persistence	1756
$\bar{P}$ Score	Single-Station Dummies	.1796
	Single-Station Boolean Pairs	.1788**
	Network Dummies	.1800
	Persistence	.1910

ACY Visibility, 7-hr Forecast (Independent Sample)

Hits	Single-Station Dummies	1774**
	Single-Station Boolean Pairs	1774**
	Network Dummies	1773
	Persistence	1674
$\bar{P}$ Score	Single-Station Dummies	.2170
	Single-Station Boolean Pairs	.2140**
	Network Dummies	.2150
	Persistence	.2270

*See Experiment No. 1.

**Asterisks are used to denote superiority.

EXPERIMENT 8. Combination of Selected Predictors from Network and Single-Station Boolean.

Purpose

To determine the extent of the improvement in the forecast when using the combined predictors selected by two separate methods.

Procedure

The 58 predictors selected by REEP in the four selection runs (CIG + 3 hrs, CIG + 7 hrs, VIS + 3 hrs, VIS + 7 hrs) using all the dummy predictors of the 13-station Atlantic City network (see Experiment No. 1) were combined with the 53 single-station predictors selected using the Boolean-paired predictors of the forecast station (see Experiment No. 7). Eliminating the duplication of predictors selected in these two ways resulted in 78 predictors, which were used in this experiment (Table 8-3). Four prediction runs were made and the results compared with the results using the network predictors (Experiment No. 1) and the Boolean-paired predictors (Experiment No. 7).

Predictands

As before, the predictands are divided into five operationally significant intervals of ceiling and visibility (Table 8-1). The specifications of the predictands are listed in Table 8-2.

Predictors

The 78 predictors used in this experiment are shown in Table 8-3, with the number of occurrences of each in the developmental sample of 7956 observations also shown. The specifications and limits of the dummy variables are as previously indicated in Table 1-4.

Data Sample

The data sample consists of 10,000 observations selected at random from the ten-year period, 1 January 1949 to 31 December 1958. The developmental sample consists of 7956 observations from the first eight years of the data sample and the verification sample consists of 2044 observations from the last two years of the ten-year period.

Results

The predictors selected by REEP for the four selection runs are shown in Tables 8-4a to 8-4d. The REEP regression equation coefficients for each group of the predictand is also shown. Contingency tables (Tables 8-5a to 8-5d) show the forecast-versus-observed frequencies for the 3-hour and 7-hour predictions of ceiling and visibility using the combined predictors as previously described. The number of correct forecasts (hits), the Brier and Allen $\bar{P}$ scores, and the individual group contributions to the overall $\bar{P}$ score are given. The comparison of the hits and the $\bar{P}$ scores achieved by these combined predictors with the hits and $\bar{P}$ scores of the network predictors alone, and the Boolean-paired single station predictors alone, is shown in Table 8-6.

Conclusions

The results of this combined study are a decided improvement over either the network or the single station Boolean predictors taken separately. This suggests that Boolean predictors generated between network variables might even improve upon the results obtained in the present combination.

Table 8-1

Definition of Predictand Groups for Atlantic City Municipal Airport

Group Number	Ceiling Intervals
1	0 ft ≤ Ceiling < 200 ft
2	200 ft ≤ Ceiling < 500 ft
3	500 ft ≤ Ceiling < 1000 ft
4	1000 ft ≤ Ceiling < 3000 ft
5	3000 ft ≤ Ceiling

Group Number	Visibility Intervals
1	0 mi ≤ Visibility < 1/2 mi
2	1/2 mi ≤ Visibility < 1 mi
3	1 mi ≤ Visibility < 2 mi
4	2 mi ≤ Visibility < 3 mi
5	3 mi ≤ Visibility

Table 8-2

Specification of Predictands - Atlantic City Municipal Airport

	Predictand 1	2	3	4
Predictand variable	CIG	CIG	VIS	VIS
Forecast interval--hours (H)	3	7	3	7
Number of Groups (G)	5	5	5	5
Observations in Group 1 (n_1)	100	135	153	167
Observations in Group 2 (n_2)	281	272	125	130
Observations in Group 3 (n_3)	385	420	200	230
Observations in Group 4 (n_4)	660	619	286	283
Observations in Group 5 (n_5)	6530	6510	7192	7146
Total dependent sample size (N)	7956	7956	7956	7956
Number of independent observations in the dependent sample (n)	2652	2652	2652	2652
Total independent sample size (M)	2044	2044	2044	2044
Number of available predictors (P)	78	78	78	78

Table 8-3

The 78 Dummy Predictors Used in Experiment 8

(See Table 1-4 for limits and specifications of each dummy)

Predictor Number	Station	Element	Dummy	Number of Occurrences
1	ACY	TOD	1	3994
2	NEL	CIG	3	411
3	NEL	TCA	2	2671
4	NEL	WEA	1	5042
5	NEL	WEA	2	423
6	NEL	WEA	9	845
7	PHL	CIG	1	42
8	PHL	CIG	5	6726
9	PHL	WND	3	1192
10	PHL	WEA	1	4760
11	PHL	WEA	7	25
12	SBY	CIG	4	654
13	SBY	CIG	5	6678
14	SBY	VIS	2	91
15	SBY	VIS	5	7438
16	SBY	WEA	2	319
17	SBY	WEA	9	751
18	MDT	VIS	2	107
19	ACY	CIG	1	109
20	ACY	CIG	2	263
21	ACY	CIG	3	410
22	ACY	CIG	4	661
23	ACY	CIG	5	6513
24	ACY	CIG	(1)	53
25	ACY	CIG	(6)	146
26	ACY	VIS	1	139
27	ACY	VIS	2	115
28	ACY	VIS	3	213
29	ACY	VIS	4	274
30	ACY	VIS	5	7215
31	ACY	RLH	3	1262
32	ACY	SCL	1	209
33	ACY	TCA	2	2633
34	ACY	WEA	1	5721
35	ACY	WEA	9	689
36	DCA	CIG	3	283
37	DCA	CIG	5	6943
38	DCA	VIS	5	7568
39	DCA	RLH	3	526
40	DCA	WEA	2	423

Table 8-3 - Continued

Predictor Number		Station	Element	Dummy	Number of Occurrences
41		DCA	WEA	9	512
42		ORF	RLH	3	791
43		PSB	CIG	1	206
44		PSB	RLH	3	1595
45		PSB	WEA	8	43
46		SWF	WND	3	1007
47		IDL	VIS	3	174
48		IDL	WND	2	1025
49		ACY	CIG	(7)	123
50		ACY	WND	4	1676
51		ACY	WND	5	2976
52		ACY	RLH	2	1418
53		ACY	WEA	2	395
54		ACY	TOD	1	841
	and	ACY	DOY	1	
55		ACY	TOD	2	258
	and	ACY	WEA	9	
56		ACY	TOD	3	243
	and	ACY	CIG	4	
57		ACY	TOD	4	118
	and	ACY	WEA	9	
58		ACY	TOD	4	290
	and	ACY	WEA	12	
59		ACY	TOD	5	505
	and	ACY	RLH	2	
60		ACY	TOD	5	425
	and	ACY	RLH	3	
61		ACY	TOD	5	160
	and	ACY	WEA	9	
62		ACY	TOD	5	105
	and	ACY	WEA	10	
63		ACY	DOY	1	578
	and	ACY	RLH	3	
64		ACY	CIG	3	383
	and	ACY	TCA	2	
65		ACY	CIG	5	6264
	and	ACY	VIS	5	
66		ACY	CIG	5	5277
	and	ACY	WEA	1	

Table 8-3 - Continued

Predictor Number	Station	Element	Dummy	Number of Occurrences
67	ACY and ACY	CIG TCA	(8) 2	332
68	ACY and ACY	WND TCA	1 2	233
69	ACY and ACY	WND TCA	2 2	496
70	ACY and ACY	WND TCA	3 1	538
71	ACY and ACY	WND TCA	3 2	754
72	ACY and ACY	WND RLH	4 3	236
73	ACY and ACY	WND TCA	4 2	492
74	ACY and ACY	WND TCA	5 2	658
75	ACY and ACY	WEA TOD	(1-8) 2	182
76	ACY and ACY	WEA DOY	(1-8) 1	215
77	ACY and ACY	WEA RLH	(1-8) 3	424
78	ACY and ACY	WEA WEA	(1-8) 9	273

Table 8-4a

Selected Predictors and REEP Equations

(Combined Network and Single Station Boolean Predictors)

Atlantic City 3-hr Ceiling

Order	Selected Predictors Station	Element	Dummy	Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1	ACY	CIG	5	.044	-.014	-.008	-.168	.146
2	ACY	CIG	1	.264	.215	.003	-.295	-.188
3	PHL	CIG	5	.008	-.021	-.066	-.129	.208
4	ACY	CIG	2	.062	.262	.009	-.216	-.118
5	ACY and ACY	TOD WEA	5 9	.164	.015	-.082	-.057	-.040
6	ACY and ACY	CIG TCA	3 2	.006	.038	.188	-.157	-.075
7	DCA	CIG	5	.004	-.021	-.050	-.050	.117
8	SBY	CIG	5	-.014	-.057	-.034	-.006	.110
9	NEL	TCA	2	.003	.007	.010	.042	-.062
10	ACY and ACY	CIG VIS	5 5	-.053	-.008	-.035	-.020	.115
11	DCA	RLH	3	.011	.067	.000	-.048	-.030
12	NEL	CIG	3	-.008	-.010	.085	-.023	-.044
13	ACY and ACY	WEA WEA	(1-8) 9	-.038	.099	-.003	-.051	-.007
14	IDL	WND	2	.009	.018	.016	.030	-.074
15	SBY	CIG	4	-.016	-.073	-.026	.098	.017
16	SWF	WND	3	.007	.023	.021	.019	-.071
17	ACY	CIG	(6)	-.007	-.016	-.041	.171	-.108
18	SBY	WEA	2	-.006	.070	-.009	.054	-.109
19	NEL	WEA	2	-.018	.003	.072	-.037	-.020
20	ACY	CIG	2	-.004	-.073	.075	-.009	.012
21	ACY and ACY	CIG TCA	(8) 2	.002	-.015	.002	.087	-.077
22	ACY	VIS	1	.069	-.052	-.063	.035	.011
23	ACY and ACY	TOD CIG	3 4	-.003	-.046	-.031	.115	-.036
	Additive Constant:			.009	.116	.190	.381	.305

Table 8-4b

Selected Predictors and REEP Equations

(Combined Network and Single Station Boolean Predictors)

Atlantic City 7-hr Ceiling

Order	Selected Predictors Station	Element	Dummy	Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1	PHL	CIG	5	.003	-.038	-.045	-.084	.164
2	DCA	CIG	5	-.001	-.001	-.040	-.074	.116
3	ACY and ACY	TOD WEA	5 9	.174	.033	-.006	-.100	-.101
4	ACY	TCA	2	-.007	-.020	.000	.052	-.026
5	IDL	WND	2	-.002	.022	.040	.035	-.096
6	PHL	WND	3	.023	.012	.025	.044	-.104
7	ACY	CIG	5	-.006	-.022	-.029	-.070	.127
8	DCA	WEA	2	-.016	.052	.117	-.026	-.127
9	MEL	WEA	9	.014	.019	.056	.011	-.100
10	ACY and ACY	WND TCA	2 2	.020	.067	.093	.012	-.192
11	ACY and ACY	WND TCA	3 2	.025	.063	.055	-.000	-.142
12	ACY and ACY	TOD WEA	4 9	.104	.017	-.029	-.076	-.017
13	ACY and ACY	TOD DOY	1 1	-.017	-.022	-.032	.066	.006
14	ACY and ACY	DOY RLH	1 3	-.005	-.038	.082	.029	-.068
15	DCA	CIG	3	.000	.088	-.003	-.053	-.032
16	ACY	CIG	2	-.008	.087	-.002	-.127	.051
17	ACY	SCL	1	.073	-.041	-.037	-.022	.027
18	SBY	WEA	9	.011	.046	.002	.030	-.088
19	PSB	RLH	3	.011	.004	.019	.022	-.056
20	ACY and ACY	WEA TOD	(1-8) 2	-.014	.010	.106	-.035	-.067
21	PSB	CIG	1	.002	.076	-.037	-.025	-.016
	Additive Constant:			.007	.064	.111	.237	.581

Table 8-4c

Selected Predictors and REEP Equations

(Combined Network and Single Station Boolean Predictors)

Atlantic City 3-hr Visibility

Order	Selected Predictors				Coefficients of the REEP Equations				
		Station	Element	Dummy	A_1	A_2	A_3	A_4	A_5
1		ACY	VIS	5	-.034	-.017	-.056	-.053	.161
2		ACY	VIS	1	.229	.037	-.023	-.158	-.085
3		ACY	RLH	3	.027	.027	.036	.023	-.114
4		ACY	VIS	2	.089	.173	-.019	-.077	-.166
5		DCA	WEA	9	.009	.010	.029	.047	-.095
6		SBY	VIS	5	-.001	-.020	-.026	-.060	.107
7		ACY	TOD	5	.122	.066	.043	-.019	-.212
	and	ACY	WEA	9					
8		ACY	TOD	5	.131	-.001	.028	-.030	-.127
	and	ACY	WEA	10					
9		NEL	WEA	9	.008	.011	.020	.052	-.091
10		ACY	WEA	(1-8)	-.068	-.023	.015	-.009	.085
	and	ACY	RLH	3					
11		ACY	WEA	1	-.001	-.003	-.008	-.045	.056
12		MDT	VIS	2	-.019	.021	.119	-.027	-.095
13		ORF	RLH	3	.035	.009	.004	.005	-.053
14		ACY	WND	4	-.005	-.031	-.072	.003	.104
	and	ACY	RLH	3					
15		ACY	CIG	1	.102	-.008	.001	-.017	-.078
16		ACY	WEA	(1-8)	.006	.006	.053	.078	-.143
	and	ACY	TOD	2					
17		ACY	TOD	4	.028	.048	.042	.023	-.141
	and	ACY	WEA	9					
			Additive Constant:		.034	.042	.093	.164	.667

Table 8-4d

Selected Predictors and REEP Equations

(Combined Network and Single Station Boolean Predictors)

Atlantic City 7-hr Visibility

Order	Selected Predictors Station		Element	Dummy	Coefficients of the REEP Equations A_1	A_2	A_3	A_4	A_5
1	and	ACY	CIG	5	-.010	-.005	-.019	-.012	.046
		ACY	VIS	5					
2	and	ACY	TOD	5	.083	.017	.046	.035	-.180
		ACY	RLH	3					
3		DCA	VIS	5	-.040	-.033	-.035	-.071	.179
4		ACY	TOD	1	-.019	-.009	-.011	-.013	.052
5		PHL	WEA	1	-.009	-.004	-.018	-.022	.054
6		ACY	WEA	9	.063	.030	.058	.022	-.173
7		DCA	WEA	2	-.026	-.018	.062	.078	-.096
8		IDL	VIS	3	.085	-.003	-.037	.017	-.062
9		PHL	WND	3	.013	.020	.020	.013	-.067
10		ACY	VIS	2	.061	.087	-.023	-.036	-.089
11	and	ACY	TOD	2	-.070	-.054	-.033	-.011	.168
		ACY	WEA	9					
12	and	ACY	TOD	5	.017	.016	.036	.042	-.111
		ACY	RLH	2					
13	and	ACY	TOD	4	.028	.024	.032	.034	-.118
		ACY	WEA	12					
14		ACY	CIG	2	-.001	.061	.019	-.020	-.059
15	and	ACY	WND	2	-.003	.019	.035	.034	-.086
		ACY	TCA	2					
			Additive Constant:		.069	.049	.076	.117	.689

Table 8-5a

Contingency Tables Showing Predicted-Versus-Actual Groups for (1) Combined Predictors Selected from Network and Single-Station Boolean, (2) Network Alone, and (3) Single-Station Boolean Alone

Atlantic City 3-hr Ceiling (Verification Sample)

REEP Using Combined Predictors

Group	Actual 1	2	3	4	5	Total
1	14	1	2	0	3	20
2	14	35	25	7	2	83
3	0	10	40	14	6	70
4	0	6	19	44	21	90
5	16	21	41	100	1603	1781
Total	44	73	127	165	1635	2044

Hits = 1736 $\quad \bar{P} = .2114$

P_g	.0155	.0276	.0422	.0600	.0660

REEP Using Network Predictors (See Experiment No. 1)

Group	Actual 1	2	3	4	5	Total
1	6	4	1	0	1	12
2	19	34	25	8	4	90
3	0	11	42	14	5	72
4	1	6	19	41	20	87
5	18	18	40	102	1605	1783
Total	44	73	127	165	1635	2044

Hits = 1728 $\quad \bar{P} = .2135$

P_g	.0163	.0277	.0423	.0605	.0666

REEP Using Single-Station Boolean Predictors (See Experiment No. 7)

Group	Actual 1	2	3	4	5	Total
1	12	3	1	0	1	17
2	19	38	27	10	10	104
3	0	9	42	10	6	67
4	1	7	16	40	29	93
5	12	16	41	105	1589	1763
Total	44	73	127	165	1635	2044

Hits = 1721 $\quad \bar{P} = .2275$

P_g	.0158	.0277	.0442	.0626	.0772

Table 8-5b

Contingency Tables Showing Predicted-Versus-Actual Groups for (1) Combined Predictors Selected from Network and Single-Station Boolean, (2) Network Alone, and (3) Single-Station Boolean Alone

Atlantic City 7-hr Ceiling (Verification Sample)

		Actual					
	Group	1	2	3	4	5	Total
	1	0	2	1	2	1	6
	2	7	12	6	0	2	27
REEP Using	3	2	18	39	14	9	82
Combined Predictors	4	5	4	14	13	4	40
	5	27	43	87	129	1603	1889
	Total	41	79	147	158	1619	2044
		Hits = 1667			$\bar{P}$ = .2576		
	P_g	.0184	.0320	.0531	.0661	.0879	

		Actual					
	Group	1	2	3	4	5	Total
	1	0	0	0	0	0	0
REEP Using	2	3	9	9	1	2	24
Network Predictors	3	7	19	34	12	9	81
	4	8	7	22	20	6	63
(See Experiment No. 1)	5	23	44	82	125	1602	1876
	Total	41	79	147	158	1619	2044
		Hits = 1665			$\bar{P}$ = .2600		
	P_g	.0180	.0323	.0540	.0666	.0891	

		Actual					
	Group	1	2	3	4	5	Total
	1	2	1	0	2	3	8
REEP Using	2	6	11	9	5	3	34
Single-Station	3	6	12	26	9	5	58
Boolean Predictors	4	4	7	13	8	6	38
(See Experiment No. 7)	5	23	48	99	134	1602	1906
	Total	41	79	147	158	1619	2044
		Hits = 1649			$\bar{P}$ = .2676		
	P_g	.0184	.0326	.0567	.0660	.0939	

Table 8-5c

Contingency Tables Showing Predicted-Versus-Actual Groups for (1) Combined Predictors Selected from Network and Single-Station Boolean, (2) Network Alone, and (3) Single-Station Boolean Alone

Atlantic City 3-hr Visibility (Verification Sample)

REEP Using Combined Predictors

	Actual					
Group	1	2	3	4	5	Total
1	16	9	3	0	2	30
2	2	2	3	0	5	12
3	0	3	0	1	1	5
4	0	1	4	3	0	8
5	24	31	68	89	1777	1989
Total	42	46	78	93	1785	2044

Hits = 1798 $\quad \bar{P} = .1772$

P_g	.0162	.0200	.0330	.0389	.0690

REEP Using Network Predictors (See Experiment No. 1)

	Actual					
Group	1	2	3	4	5	Total
1	13	9	5	1	4	32
2	2	5	4	0	9	20
3	0	2	0	0	0	2
4	0	0	3	4	0	7
5	27	30	66	88	1772	1983
Total	42	46	78	93	1785	2044

Hits = 1794 $\quad \bar{P} = .1800$

P_g	.0169	.0199	.0331	.0395	.0707

REEP Using Single-Station Boolean Predictors (See Experiment No. 7)

	Actual					
Group	1	2	3	4	5	Total
1	16	9	7	1	4	37
2	0	1	0	0	3	4
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	26	36	71	92	1778	2003
Total	42	46	78	93	1785	2044

Hits = 1795 $\quad \bar{P} = .1788$

P_g	.0165	.0201	.0327	.0389	.0705

Table 8-5d

Contingency Tables Showing Predicted-Versus-Actual Groups for (1) Combined Predictors Selected from Network and Single-Station Boolean, (2) Network Alone, and (3) Single-Station Boolean Alone

Atlantic City 7-hr Visibility (Verification Sample)

REEP Using Combined Predictors

	Actual					
Group	1	2	3	4	5	Total
1	1	0	1	1	1	4
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	42	38	89	98	1773	2040
Total	43	38	90	99	1774	2044

Hits = 1774 $\quad \bar{P} = .2141$

P_g	.0193	.0179	.0397	.0447	.0924

REEP Using Network Predictors (See Experiment No. 1)

	Actual					
Group	1	2	3	4	5	Total
1	0	0	1	0	1	2
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	43	38	89	99	1773	2042
Total	43	38	90	99	1774	2044

Hits = 1773 $\quad \bar{P} = .2150$

P_g	.0196	.0178	.0394	.0446	.0936

REEP Using Single-Station Boolean Predictors (See Experiment No. 7)

	Actual					
Group	1	2	3	4	5	Total
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	43	38	90	99	1774	2044
Total	43	38	90	99	1774	2044

Hits = 1774 $\quad \bar{P} = .2140$

P_g	.0196	.0178	.0398	.0443	.0924

Table 8-6

Comparison Between (1) Combined Predictors Selected from Network and Single-Station Boolean, (2) Network Predictors Alone*, and (3) Single-Station Boolean Predictors Alone**.

(Verification Sample)

ACY Ceiling, 3-hour Forecast

Hits	Combined Predictors	1736#
	Network Predictors	1728
	Single-Station Predictors	1721
$\bar{P}$ Score	Combined Predictors	.2114#
	Network Predictors	.2135
	Single-Station Predictors	.2275

ACY Ceiling, 7-hour Forecast

Hits	Combined Predictors	1667#
	Network Predictors	1665
	Single-Station Predictors	1649
$\bar{P}$ Score	Combined Predictors	.2576#
	Network Predictors	.2600
	Single-Station Predictors	.2676

ACY Visibility, 3-hour Forecast

Hits	Combined Predictors	1798#
	Network Predictors	1794
	Single-Station Predictors	1795
$\bar{P}$ Score	Combined Predictors	.1772#
	Network Predictors	.1800
	Single-Station Predictors	.1788

ACY Visibility, 7-hour Forecast

Hits	Combined Predictors	1774#
	Network Predictors	1773
	Single-Station Predictors	1774#
$\bar{P}$ Score	Combined Predictors	.2141
	Network Predictors	.2150
	Single-Station Predictors	.2140#

*See Experiment No. 1.
**See Experiment No. 7.
#Denotes superiority.

5.0 DISCUSSION

The REEP procedure has the following features to recommend its use over MDA:

Feature 1: The series of G regressions in REEP need not be restricted to mutually exclusive or exhaustive events as in MDA. Thus, any combination of conditions may be done simultaneously. For example, the following series of predictand elements might be considered in a single REEP analysis, where G = 15.

Y_1 = equal to unity if 0 ft $\leq$ ceiling < 200 ft

Y_2 = equal to unity if 200 ft $\leq$ ceiling < 500 ft

Y_3 = equal to unity if 500 ft $\leq$ ceiling < 1000 ft

Y_4 = equal to unity if 1000 ft $\leq$ ceiling < 1500 ft

Y_5 = equal to unity if 1500 ft $\leq$ ceiling

Y_6 = equal to unity if 0 mi $\leq$ visibility < 1/2 mi

Y_7 = equal to unity if 1/2 mi $\leq$ visibility < 1 mi

Y_8 = equal to unity if 1 mi $\leq$ visibility < 2 mi

Y_9 = equal to unity if 2 mi $\leq$ visibility < 3 mi

Y_{10} = equal to unity if 3 mi $\leq$ visibility

Y_{11} = equal to unity if 0 ft $\leq$ ceiling < 200 ft or 0 mi $\leq$ visibility < 1/2 mi

Y_{12} = equal to unity if 200 ft $\leq$ ceiling < 500 ft or 1/2 mi $\leq$ visibility < 1 mi

Y_{13} = equal to unity if 500 ft $\leq$ ceiling < 1000 ft or 1 mi $\leq$ visibility < 2 mi

Y_{14} = equal to unity if $1000 \text{ ft} \leq \text{ceiling} < 1500 \text{ ft}$ or $2 \text{ mi} \leq \text{visibility} < 3 \text{ mi}$

Y_{15} = equal to unity if $1500 \text{ ft} \leq \text{ceiling}$ or $3 \text{ mi} \leq \text{visibility}$

This feature has been made use of in Experiment 5.

Feature 2: Predictor selection in REEP is not biased by the group frequencies as is the case in MDA predictor selection. This is because each predictand variable (e.g., $Y_1, Y_2, \ldots, Y_G$) is considered separately and distinctly during selection. Most importantly, at the time of termination of selection, the significance test says there is no predictor able to provide a significant contribution to any one of the G predictand variables. For example, if the series of predictand variables is the usual five ceiling categories then predictors able to discriminate category 1 (ceiling < 100 ft) are being considered on an equal basis with those able to discriminate category 5 (ceiling $\geq$ 1500 ft). Another important feature in REEP selection is that the effective degrees of freedom, after account is taken of serial correlation, may be determined for each predictand variable separately and distinctly. For example, high ceiling conditions have a very high serial correlation (low effective degrees of freedom for nearly consecutive observations) whereas low ceiling conditions have a low serial correlation (high

effective degrees of freedom for consecutive observations). The facility to avoid the use of an average, as is unfortunately required in MDA, aids somewhat in selecting predictors having discriminatory information regarding the low conditions.

Feature 3: The mode of calculation in the REEP procedure has an additive structure in place of the multiplicative structure in MDA. That is, for both development and operational use REEP needs only the logical and arithmetic operations of addition which are faster relative to the arithmetic operations of multiplication required in the distance neighborhood portion of MDA. This has a decided effect on the speed with which both the statistical development and operational forecasting may be performed using the two techniques. The factor is in favor of REEP over MDA by about four to one (confirmed) in development and by 300 to one (estimated) in making operational forecasts.

Feature 4: It appears that REEP permits physical and synoptic interpretation of the selected predictors and the effect each has on the probability of the event being forecast. For example, in Experiment 1 (Table 1-5c) it can be seen that if the 9th predictor is a one (condition PHL ceiling < 200 ft is being observed) then .183 is added to the probability that the visibility condition three hours hence will be < 1/2 mile, whereas the other visibility categories add or subtract the following probability amounts: .018, -.079, -.084, and

-.039. However, the danger still exists for misinterpretation as a result of intercorrelation among predictors.

Feature 5: In operational forecasting, the REEP procedure can deal more efficiently than MDA with the problem of missing data or with previously unobserved weather combinations. The reason for this is that the output from REEP is a set of IBM cards containing a series of regression equations and a matrix and vector of crossproducts. If the value of a predictor variable is unavailable for making an operational forecast, a new equation omitting the missing predictor value may be generated from the output information. This operation takes very little time on a computer even when more than one predictor value is missing. MDA on the other hand has to have a matrix diagonalization performed to account for the missing predictor or predictors. In addition, discriminant functions must be calculated for the entire developmental sample of data. In REEP there is obviously no difficulty in making a prediction when the combination of predictors is not observed in the developmental sample. However, in MDA the output (on magnetic tape) is a table within which all predictor combinations observed in the developmental sample are given with the corresponding probability forecasts. When a new combination of events arises probabilities must be determined using distance neighborhood. The frequency of occurrence is about one out of every ten operational forecasts. This requirement is very

time consuming making it much less efficient than REEP in those instances.

6.0 FUTURE WORK

The most profitable single item for future effort is the modification of the prediction model so as to introduce some form of nonlinear capability. A suitable approach to this problem has already been suggested by J. G. Bryan[1]. The model consists of a linear exponential function which restricts the probability estimates to the interval (0,1). An additional feature of this model is that it uses the present REEP selection method to obtain a set of selected predictors and the REEP procedure to form a first estimate of the required coefficients. It therefore does not require an extended developmental period before it can be tested.

Another area for possible future work is a study of the selection criterion employed by REEP. The MDA criterion has been the subject of extensive research. Several possible alternatives are available and some have been looked into but none have been subjected to a fair comparison with the present method.

A combination of MDA and REEP to form a somewhat different forecast system should also be considered. This consists of applying the REEP method to the points enclosed in the Fix-Hodges neighborhood used by MDA. The purpose of this approach would be to replace the relative frequencies, now used as probabilities within the neighborhood, by the REEP estimates. While a detailed discussion of this method cannot be included here, it can be reported that this has been done for some selected cases with results which are quite encouraging.

---

[1] Bryan, J. G., 1963: Parametric estimation of probabilities in multiple group classification. Unpublished manuscript. The Travelers Research Center, Inc., 13 pp.

The matter of transforming continuous predictors to dummy predictors needs further investigation. Considerations include: means for introducing as much synoptic and/or dynamic information as possible into the dummies; experimentation with different logical combinations of the dummies; the effect of suppressing the within class variance of the continuous predictors which is an integral part of the dummying process; and the forcing of a complete set of dummies onto the selection program whenever one of the sets is selected.

In the area of application of the technique many experiments remain to be carried out to determine its characteristics on different predictands and different data. The possibility of selecting a single set of predictors applicable to several predictands distributed either in space or in time has not been thoroughly explored. Several more parameters remain to be tested for both prediction and specification. Attempts to make predictions of continuous variables by increasing the number of predictand groups is also worth investigating. In summary, a great deal of work remains to be done in the application area.

7.0 ACKNOWLEDGEMENTS

Development of the REEP procedure was substantially aided by the efforts of Mr. Thomas G. Johnson. Programming of REEP for the IBM-7090 was performed by Mr. Johnson, Miss Edna L. Sorenson, and Miss Katherine Fogarty. Mr. Karl A. Abrahamson assisted in the preparation of the report. The manuscript was typed by Mrs. Edward Miller. The author is grateful to all of these individuals.

8.0 BIBLIOGRAPHY

1. Lund, Iver A., 1955: "Estimating the Probability of a Future Event from Dichotomously Classified Predictors." Bulletin of the American Meteorological Society, Vol. 36, No. 7, pp. 325-328.

2. _____, 1961: "Some Application of the Method of Least Squares to Estimating the Probability of a Future Event." GRD Research Notes, No. 51. Geophysics Research Directorate, Air Force Cambridge Research Laboratories, Bedford, Massachusetts, 13 pp.

3. Cox, D. R., 1958: "Two Further Applications of a Model for Binary Regression." Biometrika, Vol. 45, pp. 562-565.

4. _____, 1958: "The Regression Analysis of Binary Sequences." Journal of the Royal Statistical Society, Series B, Vol. 20 pp. 215-242.

5. Warner, Stanley L., 1963: "Multivariate Regression of Dummy Variates Under Normality Assumptions." Journal of the American Statistical Association, Vol. 58, No. 304, pp. 1054-1063.

6. Miller, Robert G., 1962: "Statistical Prediction by Discriminant Analysis." Meteorological Monographs, Vol. 4, No. 25, American Meteorological Society, Boston, 54 pp.